BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers’ comments and the authors’ responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (http://bmjopen.bmj.com).

If you have any questions on BMJ Open’s open peer review process please email info.bmjopen@bmj.com
An observational cross-sectional cohorted whole population study of the association of poor broadband provision with demographic and health outcomes as determined in a single deprived multi-ethnic health economy: The Wolverhampton Digital Enablement (WODEN) Programme

<table>
<thead>
<tr>
<th>Journal:</th>
<th>BMJ Open</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>bmjopen-2022-065709</td>
</tr>
<tr>
<td>Article Type:</td>
<td>Original research</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>14-Jun-2022</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Philp, Fraser; University of Liverpool, School of Health Sciences Faux-Nightingale, Alice; Keele University, Bateman, James; Royal Wolverhampton Hospitals NHS Trust, New Cross Hospital Clark, Heather; City of Wolverhampton Council Johnson, Oliver; Point Topic Klaire, Vijay; Royal Wolverhampton Hospitals NHS Trust, New Cross Hospital Nevill, Alan; University of Wolverhampton, School of Sport Performing Arts and Leisure Parry, Emma; Keele University, Research Institute for Primary Care and Health Sciences Warren, Kate; City of Wolverhampton Council Pandyan, Anand; Bournemouth University Faculty of Health and Social Sciences Singh, Baldev; Royal Wolverhampton Hospitals NHS Trust, New Cross Hospital; University of Wolverhampton, School of Medicine &amp; Clinical Practice</td>
</tr>
<tr>
<td>Keywords:</td>
<td>EPIDEMIOLOGY, COVID-19, GENERAL MEDICINE (see Internal Medicine), HEALTH SERVICES ADMINISTRATION &amp; MANAGEMENT, Health policy &lt; HEALTH SERVICES ADMINISTRATION &amp; MANAGEMENT, Organisational development &lt; HEALTH SERVICES ADMINISTRATION &amp; MANAGEMENT</td>
</tr>
</tbody>
</table>
I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd (“BMJ”) its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge (“APC”) for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author’s Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.
An observational cross-sectional cohorted whole population study of the association of poor broadband provision with demographic and health outcomes as determined in a single deprived multi-ethnic health economy: The Wolverhampton Digital Enablement (WODEN) Programme

Authors
Fraser Philp, University of Liverpool, Liverpool, UK.
Alice Faux-Nightingale, Keele University, Newcastle under Lyme, UK.
James Bateman, Royal Wolverhampton Hospital, Wolverhampton, UK and Institute of Clinical Sciences, The University of Birmingham
Heather Clark, City of Wolverhampton Council, Wolverhampton, UK.
Oliver Johnson, Point Topic, London, UK.
Vijay Klaire, Royal Wolverhampton Hospital, Wolverhampton, UK
Alan Nevill, FEHW, University of Wolverhampton, Wolverhampton, UK
Emma Parry, School of Medicine, Keele University, Keele, UK and Royal Wolverhampton Hospital, Wolverhampton, UK.
Kate Warren, Royal Wolverhampton Hospital, Wolverhampton, UK.
Anand Pandyan, Bournemouth University, Bournemouth, UK
Baldev Singh, Royal Wolverhampton Hospital, Wolverhampton, UK and University of Wolverhampton, Wolverhampton, UK.

Keywords
HEALTH SERVICES ADMINISTRATION & MANAGEMENT, EPIDEMIOLOGY, COVID-19, GENERAL MEDICINE (see internal medicine), Health policy < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Organisational development < HEALTH SERVICES ADMINISTRATION & MANAGEMENT
Corresponding Author

Professor Baldev M SINGH, The Royal Wolverhampton NHS Trust, New Cross Hospital, Wolverhampton, WV10 0QP, UK & School of Medicine and Clinical Practice, Faculty of Science and Engineering, University of Wolverhampton, Wolverhampton, WV1 1LY, UK; email: baldev.singh@nhs.net, https://orcid.org/0000-0003-3460-6759

Conflict of Interests

None

Copyright

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence on a worldwide basis to the BMJ Publishing Group Ltd to permit this article (if accepted) to be published in BMJ editions and any other BMJPG products and sublicenses such use and exploit all subsidiary rights, as set out in our licence.

Word count

3145
Abstract

Objectives

To determine the association between measurable broadband provision and demographic and health outcomes in a defined population.

Design

An observational cross-sectional whole local population level study with cohorts defined according to broadband provision.

Setting / Participants

Data for all residents of the City of Wolverhampton, totalling 269,785 residents.

Primary outcomes

Poor broadband provision is associated with variation in demographics and with increased comorbidity and urgent care needs.

Results

Broadband provision was measured using the Broadband Infrastructure Index (BII) in 158 City localities housing a total of 269,785 residents. Lower broadband provision as determined by BII was associated with younger age (p<0.001), white ethnic status (p<0.001), lesser deprivation as measured by Index of Multiple
Deprivation (p<0.001), a higher number of health comorbidities (p<0.001) and more non-elective urgent events over 12 months (p<0.001).

**Conclusion**

Local municipal and health authorities are advised to consider the variations in broadband provision within their locality and determine equal distribution both on a geographical basis but also against demographic, health and social data to determine equitable distribution as a platform for equitable access to digital resources for their residents.

**Strengths and limitations of this study**

- This is the first study to link data on broadband provision to health data at a defined health economy population level.
- Our observational study used a high-quality dataset from the local population that has been used in other research studies.
- Our study methodology, which is generalisable, enables assessment of the importance of evaluation at the local level considering local population variations, perhaps lost at the inter-regional or national level.
- The study is limited by being observational and cross-sectional rather than interventional or prospective.
- The study does not have individual-level socio-economic data, nor an individual assessment of digital access, use or competency.
Introduction

The very rapid integration of digital technology into the delivery of healthcare has been accelerated by Covid-19\(^1,2\). Driven by the pandemic, the adaptability of healthcare providers to replace “face to face” consultations with technology-based remote access services has been truly commendable. However, such digital services may not have been accessible to all service users. Clearly recognised access barriers are a user’s ability to engage with technology and their digital illiteracy\(^2\) but others include limited access to any required equipment, broadband connectivity and internet access. Such infrastructure considerations may well be equally fundamental to digital healthcare service access. All such factors are broadly recognised as the digital determinants of health\(^1\).

Since the digital transformation of health care is inevitable, there is a need to ensure that service users are thereby not excluded. Identifying those with digitally driven access barriers, and ensuring that services are flexed to provide equity of care to them, is important both to avoid direct clinical risk to their management and to prevent a widening of societal health inequalities\(^3,4,5-8\). Such associations have raised the spectre of creating a digital underclass\(^5,6,9\).

Accordingly, the play of the digital determinants at the individual or group level are increasingly being considered\(^10-12\) but that does not seem to have led to a clear or systematic approach at the level of health care economies. Healthcare authorities will increasingly need to develop a more comprehensive understanding of population level needs and shift resources accordingly. Even then, it is not clear that we have the evidence-based assessments to inform such decision making and such tools that exist are generally applied to individuals. This evidence is usually gathered through questionnaires that have variable validity and are limited by small sample sizes or are conducted on discrete groups\(^13\), often those with standing digital access.

Wolverhampton is a UK City functioning as a single health and social care economy with a population of circa 270,000. It is multi-ethnic and socio-economically deprived\(^14\). The Wolverhampton Digital Enablement (WODEN) Programme is a multi-agency collaborative approach to determine and address digital factors that may impact on health and social care. The aim of this study was to focus on broadband provision and to consider, at population scale, its association with demography and with key health outcomes. Published research that systematically analyses the association between impaired digital access with health outcomes is very limited. As far we can determine, our methodology is novel, and any independent association with adverse health outcome, whilst intuitively expected, has never been systematically determined.
Methods

Population studied

All patients registered with a Wolverhampton General Practitioner (GP) and those who were residents known to have had any hospital contact.

Data

Individual data were anonymised. To identify the link between broadband provision, demographic and health outcomes, a single dataset was developed by merging the Point Topic Broadband Digital Exclusion\textsuperscript{15} and the Wolverhampton Integrated Health Care databases using local authority area codes for the Lower Layer Super Output Areas (LLSOAs). LLSOAs describes a geographical breakdown of the UK, with an average population of 1500 people or 650 households, and are widely used to improve the reporting of small populations in NHS data modelling.

The Wolverhampton Integrated Healthcare database

The Wolverhampton integrated healthcare database is a database developed to integrate data from primary care, community and hospital clinical and pathology systems, for the residents of Wolverhampton, and from immediately adjacent districts, for all people residing in Wolverhampton or registered to Wolverhampton practices, and from immediately adjacent districts with emergency admission to New Cross Hospital (NXH), using methodology we have previously described\textsuperscript{16}.

The demographic and health care variables extracted were age, gender, ethnicity, the Index of Multiple Deprivation (IMD) ranked score as determined from postcodes, long term condition comorbidities and non-elective health activity. Ethnicity data from all sources were reviewed, only unambiguous data were accepted, then recoded into White, South-Asian, Black, Mixed Ethnicity, Chinese or Unknown. The sum of co-morbidities in any individual was determined by the total number of comorbidities which matched the long-term conditions found in the population which are utilised in defining co morbidity in the local health economy\textsuperscript{16} commonest comorbidities. These are: asthma, atrial fibrillation, cancer, cardiovascular diseases, chronic kidney disease, chronic obstructive pulmonary disease, dementia, depression and other mental health
disorder, diabetes, epilepsy, haemoglobinopathy, heart failure, hypertension, learning difficulties, osteoarthritis and rheumatoid arthritis. The sum of non-elective health activity events was determined by the number of emergency department attendances not leading to admission plus the number of non-elective admissions over the preceding 12 months to our local hospital.

The Point Topic database and the identification of appropriate measure of broadband provision

The Point Topic® system develops a global Broadband Digital Deprivation Index from a variety of variables including the Broadband Infrastructure Index (BII) but also the IMD summary score and other ranked measures derived from the IMD such as age, income and housing. The index is a ranked score for all English LLSOAs. The BII subcomponent was the variable selected as a measure of broadband infrastructure provision. The BII is an index measure representative of broadband provision including the type, speed and bandwidth quality and number of suppliers for broadband in those defined localities\textsuperscript{15, 17}. The BII assessment used is for the year 2020. Thus, within the merged dataset, there were a number of variables that might have shared common subcomponents and particular care was taken to ensure the independence of BII as a variable (presented in Results). We emphasise the BII to be a measure of provision, not access or uptake and, just as standard practice with the use of IMD, it defines the characteristic of a geographical unit, the LLSOA, not the individuals who may live in it.

Statistical method

All data were analysed on IBM SPPS version 26. When comparing independent groups, the Students t-test and the Chi-square test were used for the difference between means and proportions respectively. Principle components analysis (with rotation) was utilised when considering the inter-dependence of independent variables especially in relationship to the BII.

The BII, an ordinal variable, was not normally distributed and clearly dichotomised into two groups above and below a rank value of 10,000 (Figure 1). The purpose of this analysis was to define associations with the poorer provisioned broadband group with a BII of <10,000. Thus, binary logistic regression analysis was used to examine the relationship between BII as the dependent (dichotomised) variable and various independent variables. The independent variables were in two broad groupings 1) the demographic variables of age,
gender, ethnicity, and 2) variables of deprivation; the two health outcome variables of long-term condition multimorbidity and of emergency non-elective urgent care. There is a known association and thus a potential confounding effect between these two independent variable groupings and this was negated in our analysis by taking a stepwise approach, entering each block separately as previously described. The regression analysis was undertaken on 250,609 residents as 19,176 were excluded by missing or unclassifiable ethnicity coding. Results are presented as the mean ± SD or as percentages. Statistical significance was taken at p<0.05.

Ethical Approval

Not applicable.

As the data were accrued and analysed to determine the service imperative to be aware of digital exclusion as a potential factor in delivering equitable social and health care, and as no selection, or randomisation or intervention occurred, research ethical approval was not deemed necessary, and this was confirmed within local governance processes.

Patient and Public Involvement

None as not applicable to this type of study.
Results

Population and Localities

The final cohort were the 269,785 alive at the point of data acquisition in the year 2020 and included 13,945 residents registered to a non-Wolverhampton GP (Table 1). The datasets were linked on locality codes. There are 158 LLSOAs in Wolverhampton with a surface area of (mean ±SD) (range) 0.44 ±0.28 (0.13 – 2.20) km², with 691 ±139 (466 – 1388) and 1708±435 (857 – 3476) households and individuals per locality respectively. We again emphasize that IMD and BII data are derived from rankings relating to the locality codes and not from direct assessment of individuals and we do not have individual socio-economic data.

The Broadband Infrastructure Index (BII)

The wider Broadband Digital Deprivation Index is derived from many variables including those related to deprivation. Thus, the correlation between IMD (as IMD rank score) and the global Broadband Digital Deprivation Index was tight and highly significant (r = 0.902, p<0.001). In a principal component analysis, including all subsidiary variables, four components emerged (Eigen value of >=1), the BII sub-component of the Broadband Digital Deprivation Index was independent of all other IMD indices apart from IMD housing rankings, which equally spanned the deprivation and BII components, which has face validity. Accordingly, the linear correlation between BII and IMD ranking was extremely weak (r = -0.16, p<0.001) with an $r^2$ of only 3%, confirming with high confidence that the two measures were independent. The BII component was therefore identified as an appropriate measure of broadband provision for linkage to health data at the LLSOA local authority area code levels. We emphasise BII not to be a continuous variable since the group was clearly divisible into two cohorts of those with a distinctly low BII versus others at a threshold rank score of 10,000 (Figure 1).

The association between broadband access, demographic and health outcomes

The demographic and health data of the whole population and of the BII groupings are shown in Table 1. Regarding demography, the lower scoring BII group (lesser access) were older and also had marginal more females. Of particular note, for the lesser provisioned BII category, both the IMD parameters indicated this
group to be less, not more, deprived and they had a higher, not lower, white ethnicity prevalence. The crude relationship of the BII as a dichotomised group to quintiles of the IMD score ($\chi^2 = 7559.4, p<0.001$), quintiles of age ($\chi^2 = 156.7, p<0.001$) and with ethnicity ($\chi^2 = 1521.5, p<0.01$) is depicted in Figure 2. For parameters of health outcomes (Table 1), the lower scoring BII group was significantly more comorbid as measured by the total number of comorbidities or by the proportion with any comorbidity and they had a very small but significant excess of non-elective urgent health activity over the preceding 12 months. Since BII is not a continuous variable, these factors were also entered into binary logistic regression of the dichotomised BII status as the dependent variable with the purpose of highlighting the magnitude, direction and significance of any association with the lower BII group of the defined independent variables which were entered in two stages: firstly, the demographic factors of quintiles of age, gender (male vs female), ethnicity (white vs other), and quintiles of the IMD score; secondly, comorbidity (any versus none) and urgent events (any versus none).

Table 2 shows the outputs of this model which was highly significant ($\chi^2 = 7914.8, p<0.001$), albeit the independent factors explained very little of the variance ($r^2 = 0.042$). At step 1, the model adjusted Odds Ratios confirmed the crude associations with IMD ($p<0.001$) and ethnicity ($p<0.001$) but exposed a different (model adjusted) pattern for age ($p<0.001$) with the younger quintiles to be in the low scoring BII group whilst the association of gender was not significant. At step 2, the relationship of health indices of Any comorbidity ($p<0.001$) and any Urgent care event ($p<0.001$) to low BII were both highly significant, noting these two latter variables to have a degree of correlation ($\chi^2 = 3719.4, \text{Eta} 0.12, p<0.001$).
Table 1: Demographic and health data of the whole resident population, subdivided by categories of the Broadband Infrastructure Index (BII).

Results are the mean ± SD with (range), ns = non-significant and the p-value is for the significance of the difference between the two BII categories.

<table>
<thead>
<tr>
<th></th>
<th>Whole population n= 269,785</th>
<th>BII &lt; 10,000 (worse) n=125,007 (46%)</th>
<th>BII&gt;10,000 n=144,778 (54%)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BII rank</td>
<td>11,657±6,926 (3,451 – 5,095)</td>
<td>4,971±348 (2,804 – 5,095)</td>
<td>17,430±4168 (11,328 – 27,276)</td>
<td>t=1,053.6, p&lt;0.001</td>
</tr>
<tr>
<td>Age (years)</td>
<td>38.6±23.2 (0 – 106)</td>
<td>39.1±23.3</td>
<td>38.3±23.0</td>
<td>t=8.816, p&lt;0.001</td>
</tr>
<tr>
<td>Gender (male %)</td>
<td>50.4%</td>
<td>50.1%</td>
<td>50.6%</td>
<td>χ²=4.45, p&lt;0.05</td>
</tr>
<tr>
<td>Ethnicity White %</td>
<td>57.1%</td>
<td>65.6%</td>
<td>57.9%</td>
<td>χ²=1521.5, p&lt;0.001</td>
</tr>
<tr>
<td>IMD Score</td>
<td>34.7±15.9 (5.4 – 71.8)</td>
<td>32.6±15.6</td>
<td>36.4±15.8</td>
<td>t=63.549, p&lt;0.001</td>
</tr>
<tr>
<td>IMD Rank</td>
<td>8878±7872</td>
<td>9698±7739</td>
<td>8169±7916</td>
<td>t=50.586, p&lt;0.001</td>
</tr>
<tr>
<td>Number of comorbidities</td>
<td>0.61±1.06 (0 – 11)</td>
<td>0.63±1.07</td>
<td>0.59±1.05</td>
<td>t=11.069, p&lt;0.001</td>
</tr>
<tr>
<td>Any comorbidity</td>
<td>35%</td>
<td>36.5%</td>
<td>34%</td>
<td>χ²=173.54, p&lt;0.001</td>
</tr>
<tr>
<td>Number of non-elective contacts in 12 months</td>
<td>0.24±0.84 (0 – 65)</td>
<td>0.242±0.85</td>
<td>0.235±0.83</td>
<td>t=2.042, p&lt;0.05</td>
</tr>
<tr>
<td>Any non-elective contact in 12 months</td>
<td>14.4 %</td>
<td>14.7%</td>
<td>14.4%</td>
<td>χ²=16.01, p&lt;0.001</td>
</tr>
</tbody>
</table>
Table 2: Binary logistic regression identifying factors associated with the poor or least provisioned Broadband Infrastructure Index (BII) group.

The analysis was undertaken on 250,609 subjects with 19,176 missing cases due to a lack of ethnicity data. The Odds ratios (OR) are for the poor or least provisioned BII group. IMD is the Index of Multiple Deprivation. An * represents the reference group for each categorical variable from which all other groups in that variable were compared.

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR (95% CI)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMD Q1 (11.6±3.2) (least deprived)</td>
<td>2.14 (2.09-2.19)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>IMD Q2 (24.3±4.0)</td>
<td>2.37 (2.31-2.43)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>IMD Q3 (36.8±2.5)</td>
<td>1.38 (1.35-1.42)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>IMD Q4 (44.1±2.0)</td>
<td>1.27 (1.24-1.31)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>IMD Q5 (56.1±5.2) (most deprived)</td>
<td>1.0</td>
<td>-</td>
</tr>
<tr>
<td>Age Q1 (7.7±4.5 years) (youngest)</td>
<td>1.20 (1.16-1.23)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age Q2 (24.0±4.7 years)</td>
<td>1.08 (1.05-1.11)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age Q3 (38.3±4.0 years)</td>
<td>1.08 (1.05-1.11)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age Q4 (52.8±4.3 years)</td>
<td>1.06 (1.03-1.09)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age Q5 (72.8±8.7 years) (oldest) *</td>
<td>1.0</td>
<td>-</td>
</tr>
<tr>
<td>White Ethnicity (61.5%)</td>
<td>1.33 (1.31-1.36)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Other ethnic groups *</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td><strong>Step 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any comorbidities present (35.2%)</td>
<td>1.07 (1.05-1.09)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No comorbidities*</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Any emergency episode (14.4%)</td>
<td>1.05 (1.03-1.08)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No emergency episodes*</td>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1 Demonstrating the dichotomised distribution of the Broadband Infrastructure Index (BII) and its suitability for binary logistic regression analysis.

Figure 2 The crude relationships of IMD, age and Ethnicity to the category of the Broadband Infrastructure Index (BII).
Discussion

On the contention that physical availability of broadband infrastructure should not be overlooked as a precursor to health care access\textsuperscript{19}, we aimed to identify the nature of association between broadband provision, population demographics and health outcomes in a single health economy, ours being deprived and multi-ethnic by UK standards. To our knowledge, this is the first study to identify a candidate measure of broadband provision (BII) and link it to population-level health data (urgent care and comorbidity) and indices of deprivation (IMD). In some sense, we wished to consider the epidemiology of broadband provision. We find that associations exist between BII determined broadband provision, demographic variables and health outcomes.

The derivation of the BII is described elsewhere\textsuperscript{15, 17}. Previous studies, to the best of our knowledge, have not identified or utilised any adequate measure of broadband infrastructure\textsuperscript{10, 20-22} and we are the first to demonstrate that it is essentially an independent factor so long as it is disaggregated from association with deprivation data when used in this context. Had we used the broader more global Broadband Digital Deprivation Index as the measure, these associations would not have been seen, confounded by strong collinearity with IMD. BII in its own right can be utilised to demonstrate variation by geographical locality. Our own data showing localities that stood out as less well provisioned emphasises the point that any such avoidable digital divide is an addressable infrastructure issue. How such variation arose is unknown to us, perhaps by historical chance events, or other practical infrastructure constraints, or because broadband providers are driven by commercial and fiscal but not social imperatives\textsuperscript{21}. Nevertheless, geographical variation of provision can only be addressed if measured.

We demonstrate in the crude data and in regression models that BII was associated with variation with demography and with measurable health outcomes. In binary logistic regression, contrary to expectation, the lower provisioned BII group were less deprived, had a lower non-white ethnicity make-up, and the adjusted association with age showed an impact in younger age groups. The association of poorer broadband provision and the health outcomes of both co-morbidity and non-elective urgent care has not been previously described at a population level. There arise a number of considerations: whilst the associations were highly statistically significant the size of the effect was very small with the independent variables describing < 5% of the variance of BII. We are thus simply stating that they can exist. We believe our local findings are unlikely to be definitive and will vary from region to region, the point being that, unless measured, accountable organisations cannot
be assured of equitable provision as a platform for equitable access other than perhaps by geography. Regarding geography, published comparisons are generally made between large population blocks at regional or national level\textsuperscript{14, 17, 23}, whereas we have considered broadband provision within a single local healthcare economy drilling down to the LLSOAs\textsuperscript{17} and determined that focus at that level ought to be a consideration for local municipal and health authorities.

The distinction between provision and uptake also requires emphasis, perhaps crudely summarised as the distinction between public provision and individual access. The latest UK government and other independent analysis shows rapidly increasing internet usage\textsuperscript{14, 17, 23} with less than 10% of the population being non or infrequent users, although poor usage remains associated with lower income, lesser educational attainment, increasing age, to a much lesser extent ethnicity and gender as well as geographical factors including the rural urban divide. At an individual level, the skill set required for internet usage and the nature of that use is increasingly understood\textsuperscript{23}, noting that internet use for health remains relatively low. Regarding the skill set required, 8% of the UK population have no basic level skills and in 21% they are limited\textsuperscript{24}. Generally, there is evidence that broadband and internet usage are associated with a variety of benefits amongst those who are digitally included\textsuperscript{24, 25}. The concept of the digital divide and digital exclusion arises out of a concern for the capability among individuals and communities to benefit from digital innovations. In health care, specific measures have been developed to assess digital literacy including e-Heals\textsuperscript{26} and applied in differing digital health interventions. From that has arisen an understanding of the barriers to engagement with such digital health interventions\textsuperscript{27}, how and whether health behaviour can be modified\textsuperscript{28} and, importantly, that digital health interventions, of themselves, can lead to inequality\textsuperscript{29}.

Publicly accessible digital facilities and personal smartphones may be solutions to inequitable home broadband provision, although they have limitations\textsuperscript{17}. Broadband access within an individual’s home, rather than a community space, is thought important for increasing digital engagement and associated with higher satisfaction, a wider freedom of use and more complex activities including those related to healthcare\textsuperscript{30-32}. However, bridging the gap from infrastructure to broadband access and internet use, end point access, previously often described as ‘last mile access’, is likely determined by uptake. As broadband infrastructure provision increasingly becomes a politically recognised social determinant of health and other outcomes\textsuperscript{12, 33, 34-36}, with reports describing broadband infrastructure as a “super social determinant of health”\textsuperscript{20, 37}, it perhaps will come to be recognised as a basic social right, especially if there is an overlap between broadband provision and social inequality\textsuperscript{1, 9, 38}. If broadband provision is an investment in the health and wealth of a
For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

population, then digital poverty might require greater recognition, for which a direct analogy exits for the provision of energy utilities, fuel poverty and fuel payments.  

Caveats, Strengths and Weaknesses

We recognise that the BII is a geographical parameter and does not tell us about individual digital exclusion, since we do not have such assessments, just as the use of IMD is taken in lieu of individual detailed socio-economic data.

Our findings are pertinent to our local health economy and may well differ from similarly sized unitary health and social care geopolitical areas.

This study is a cross-sectional time point view in rapidly changing digital environment.

As a further caveat we emphasise that only a prospective follow up study can determine whether broadband provision is able to modify health determinants and outcomes, which surely must be the intent of all digital health initiatives.

With those caveats and weaknesses, we believe this to be the first description of the association between measurable broadband provision and health data at a population level and we are the first to demonstrate a statistically significant link with certain crucial health outcomes - comorbidity and urgent care.

Our study methodology is generalisable and based on a high quality and complete dataset curated for the local population. This enables assessment at the local level, taking into account local variations and moving away from large scale inter-regional or national level comparisons.

Conclusion
The provision of broadband is but one component of the digital infrastructure capacity and capability required for an individual to access and reap the potential benefits of the digitalisation health and social care. We describe a methodology to measure such provision and demonstrate that its variation can be determined at a very local level within unitary and health and social care economies. Such differences may be associated with an unintended variation in a populations’ demography and with important health imperatives. Such associations cannot be assumed but are best measured,

Implications for clinicians and policymakers

Local authorities may wish to consider and extend our described methodology to ensure equitable provision of broadband infrastructure.

Individual care providing services and their staff must be aware of the potential constraints faced by their clients and patients as digital interventions are developed as well as their own limitations.

Future research

As is the intent of the WODEN group, hopefully, our work will promote future research in this field to address the many aspects needed to ensure equity in health and care delivery, and the demonstrable avoidance of imparting disequity, as we so rapidly move into the digital paradigm.

Author contributions

Accountable senior author BMS; VK; Manuscript writing: FP, AFN, VK, KW, EP, HC, JB, OJ, AN, ADP, BMS; Preparation for submission: FP, AFN, AN, ADP, BMS; Database quality, data integration and data quality and integration: BMS, VK; Statistical advice: AN, BMS, All authors contributed intellectual content during the drafting and revision of the work and approved the final version.

Transparency declaration
BMS affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

**Funding**

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interest statement**

All authors have completed the Unified Competing Interest form (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years, no other relationships or activities that could appear to have influenced the submitted work.

**Data Sharing**

Anonymised data will be shared on reasonable request to the corresponding author.

**Ethics statement**

Not applicable.

As the data were accrued and analysed to determine the service imperative to be aware of digital exclusion as a potential factor in delivering equitable social and health care, and as no selection, or randomisation or intervention occurred, research ethical approval was not deemed necessary, and this was confirmed within local governance processes.
References


34. Ofcom. Affordability of communications services A summary of initial findings. 2020.


Figure
The crude relationships of IMD, Age and Ethnicity to the category of the Broadband Infrastructure Index (BII)

533x755mm (118 x 118 DPI)
STROBE Statement—checklist of items that should be included in reports of observational studies

<table>
<thead>
<tr>
<th>Item No.</th>
<th>Recommendation</th>
<th>Page No.</th>
<th>Relevant text from manuscript</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title and abstract</strong></td>
<td>1</td>
<td>(a) Indicate the study’s design with a commonly used term in the title or the abstract</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(b) Provide in the abstract an informative and balanced summary of what was done and what was found</td>
<td>3</td>
</tr>
<tr>
<td><strong>Introduction</strong></td>
<td>2</td>
<td>Explain the scientific background and rationale for the investigation being reported</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>State specific objectives, including any prespecified hypotheses</td>
<td>5</td>
</tr>
</tbody>
</table>
### Methods

| Study design | 4 | Present key elements of study design early in the paper |
| Setting | 5 | Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection |
| Participants | 6 | (a) Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up  
Case-control study—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  
Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants |
| (b) Cohort study—For matched studies, give matching criteria and number of exposed and unexposed  
Case-control study—For matched studies, give matching criteria and the number of controls per case |
| Variables | 7 | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable |
| 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 |
| Bias | 9 | Describe any efforts to address potential sources of bias |

Only unambiguous data were accepted.

There were a number of variables that might have shared common subcomponents and particular care was taken to ensure the independence of BII as a variable (presented in Results). We emphasise the BII to be a measure of provision,
not access or uptake and, just as standard practice with the use of IMD, it defines the characteristic of a geographical unit, the LLSOA, not the individuals who may live in it.

There is a known association and thus a potential confounding effect between these two independent variable groupings and this was negated in our analysis by taking a stepwise approach, entering each block separately as previously described.

<table>
<thead>
<tr>
<th>Study size</th>
<th>Explained how the study size was arrived at</th>
<th>6,981</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Explain how the study size was arrived at</td>
<td></td>
</tr>
</tbody>
</table>

All patients registered with a Wolverhampton General Practitioner (GP) and those who were residents known to have had any hospital contact.

The final cohort were the 269,785 alive at the point of data acquisition in the year 2020 and included 13,945 residents registered to a non-Wolverhampton GP (Table 1).
### Results

**Participants**
- (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
- (b) Give reasons for non-participation at each stage
- (c) Consider use of a flow diagram

**Descriptive data**
- (a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
- (b) Indicate number of participants with missing data for each variable of interest
- (c) **Cohort study**—Summarise follow-up time (eg, average and total amount)

**Outcome data**
- **Cohort study**—Report numbers of outcome events or summary measures over time
- **Case-control study**—Report numbers in each exposure category, or summary measures of exposure
- **Cross-sectional study**—Report numbers of outcome events or summary measures

**Main results**
- (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
- (b) Report category boundaries when continuous variables were categorized
- (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period

The regression analysis was undertaken on 250,609 residents as 19,176 were exclude by missing or unclassifiable ethnicity coding.
Other analyses  17  Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses

Discussion

Key results  18  Summarise key results with reference to study objectives

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

Interpretation  20  Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence

Generalisability  21  Discuss the generalisability (external validity) of the study results

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

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

An observational cross-sectional study of the association of poor broadband provision with demographic and health outcomes: The Wolverhampton Digital ENablement (WODEN) Programme

<table>
<thead>
<tr>
<th>Journal:</th>
<th>BMJ Open</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>bmjopen-2022-065709.R1</td>
</tr>
<tr>
<td>Article Type:</td>
<td>Original research</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>16-Aug-2022</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Philp, Fraser; University of Liverpool, School of Health Sciences Faux-Nightingale, Alice; Keele University, Bateman, James; Royal Wolverhampton Hospitals NHS Trust, New Cross Hospital Clark, Heather; City of Wolverhampton Council Johnson, Oliver; Point Topic Klaire, Vijay; Royal Wolverhampton Hospitals NHS Trust, New Cross Hospital Nevill, Alan; University of Wolverhampton, School of Sport Performing Arts and Leisure Parry, Emma; Keele University, Research Institute for Primary Care and Health Sciences Warren, Kate; City of Wolverhampton Council Pandyan, Anand; Bournemouth University Singh, Baldev; Royal Wolverhampton Hospitals NHS Trust, New Cross Hospital; University of Wolverhampton, School of Medicine &amp; Clinical Practice</td>
</tr>
<tr>
<td>Primary Subject Heading:</td>
<td>Epidemiology</td>
</tr>
<tr>
<td>Secondary Subject Heading:</td>
<td>Evidence based practice, Health policy, Health services research, Patient-centred medicine, Public health</td>
</tr>
<tr>
<td>Keywords:</td>
<td>EPIDEMIOLOGY, COVID-19, GENERAL MEDICINE (see Internal Medicine), HEALTH SERVICES ADMINISTRATION &amp; MANAGEMENT, Health policy &lt; HEALTH SERVICES ADMINISTRATION &amp; MANAGEMENT, Organisational development &lt; HEALTH SERVICES ADMINISTRATION &amp; MANAGEMENT</td>
</tr>
</tbody>
</table>
I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd (“BMJ”) its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge (“APC”) for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author’s Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.
Title page

Title

An observational cross-sectional study of the association of poor broadband provision with demographic and health outcomes: The Wolverhampton Digital ENablement (WODEN) Programme

Authors

Fraser Philp, University of Liverpool, Liverpool, UK.

Alice Faux-Nightingale, Keele University, Newcastle under Lyme, UK.

James Bateman, Royal Wolverhampton Hospital, Wolverhampton, UK and Institute of Clinical Sciences, The University of Birmingham

Heather Clark, City of Wolverhampton Council, Wolverhampton, UK.

Oliver Johnson, Point Topic, London, UK.

Vijay Klaire, Royal Wolverhampton Hospital, Wolverhampton, UK

Alan Nevill, FEHW, University of Wolverhampton, Wolverhampton, UK

Emma Parry, School of Medicine, Keele University, Keele, UK and Royal Wolverhampton Hospital, Wolverhampton, UK.

Kate Warren, Royal Wolverhampton Hospital, Wolverhampton, UK.

Anand Pandyan, Bournemouth University, Bournemouth, UK

Baldev Singh, Royal Wolverhampton Hospital, Wolverhampton, UK and University of Wolverhampton, Wolverhampton, UK.

Keywords

HEALTH SERVICES ADMINISTRATION & MANAGEMENT, EPIDEMIOLOGY, COVID-19, GENERAL MEDICINE (see internal medicine), Health policy < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Organisational development < HEALTH SERVICES ADMINISTRATION & MANAGEMENT
Corresponding Author

Professor Baldev M SINGH, The Royal Wolverhampton NHS Trust, New Cross Hospital, Wolverhampton, WV10 0QP, UK & School of Medicine and Clinical Practice, Faculty of Science and Engineering, University of Wolverhampton, Wolverhampton, WV1 1LY, UK; email: baldev.singh@nhs.net, https://orcid.org/0000-0003-3460-6759

Conflict of Interests

None

Copyright

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence on a worldwide basis to the BMJ Publishing Group Ltd to permit this article (if accepted) to be published in BMJ editions and any other BMJPGL products and sublicenses such use and exploit all subsidiary rights, as set out in our licence.

Word count

3234
Abstract

Objectives

The association between impaired digital provision, access and health outcomes has not been systematically studied. The Wolverhampton Digital ENablement (WODEN) Programme is a multi-agency collaborative approach to determine and address digital factors that may impact on health and social care in a single deprived multi ethnic health economy. The objective of this study is to determine the association between measurable broadband provision and demographic and health outcomes in a defined population.

Design

An observational cross-sectional whole local population level study with cohorts defined according to broadband provision.

Setting / Participants

Data for all residents of the City of Wolverhampton, totalling 269,785 residents.

Primary outcomes

Poor broadband provision is associated with variation in demographics and with increased comorbidity and urgent care needs.

Results

Broadband provision was measured using the Broadband Infrastructure Index (BII) in 158 City localities housing a total of 269,785 residents. Lower broadband provision as determined by BII was associated with
younger age (p<0.001), white ethnic status (p<0.001), lesser deprivation as measured by Index of Multiple
Deprivation (p<0.001), a higher number of health comorbidities (p<0.001) and more non-elective urgent
events over 12 months (p<0.001).

Conclusion

Local municipal and health authorities are advised to consider the variations in broadband provision within
their locality and determine equal distribution both on a geographical basis but also against demographic,
health and social data to determine equitable distribution as a platform for equitable access to digital
resources for their residents.

Strengths and limitations of this study

- This is the first study to link data on broadband provision to health data at a defined health economy
  population level.
- Our observational study used a high-quality dataset from the local population that has been used
  in other research studies
- Our study methodology, which is generalisable, enables assessment of the importance of
  evaluation at the local level considering local population variations, perhaps lost at the inter-
  regional or national level.
- The study is limited by being observational and cross-sectional rather than interventional or
  prospective.
- The study does not have individual-level socio-economic data, nor an individual assessment of
digital access, use or competency.
Introduction

The very rapid integration of digital technology into the delivery of healthcare has been accelerated by Covid-19\(^1,2\). Driven by the pandemic, the adaptability of healthcare providers to replace “face to face” consultations with technology-based remote access services has been truly commendable. However, such digital services may not have been accessible to all service users. Clearly recognised access barriers are a user’s ability to engage with technology and their digital illiteracy\(^2\) but others include limited access to any required equipment, broadband connectivity and internet access. Such infrastructure considerations may well be equally fundamental to digital healthcare service access. All such factors are broadly recognised as the digital determinants of health\(^1\).

There is a need to ensure that service users are not excluded as a result of historical, current and future strategies which will drive digital transformations to health care\(^3-5\). Identifying those with digitally driven access barriers, and ensuring that services are flexed to provide equity of care to them, is important both to avoid direct clinical risk to their management and to prevent a widening of societal health inequalities\(^6,7\)\(^8-11\). Such associations have raised the spectre of creating a digital underclass\(^8,9,12\).

Accordingly, the play of the digital determinants at the individual or group level are increasingly being considered\(^4,13,14\) but that does not seem to have led to a clear or systematic approach at the level of health care economies. Healthcare authorities will increasingly need to develop a more comprehensive understanding of population level needs and shift resources accordingly. Even then, it is not clear that we have the evidence-based assessments to inform such decision making and such tools that exist are generally applied to individuals. This evidence is usually gathered through questionnaires that have variable validity and are limited by small sample sizes or are conducted on discrete groups\(^15\), often those with standing digital access.

Wolverhampton is a UK City functioning as a single health and social care economy with a population of circa 270,000. It is multi-ethnic and socio-economically deprived\(^16\). The Wolverhampton Digital Enablement (WODEN) Programme is a multi-agency collaborative approach to determine and address digital factors that may impact on health and social care. The aim of this study was to focus on broadband provision and to consider, at population scale, its association with demography and with key health outcomes. Published research that systematically analyses the association between impaired digital access with health outcomes is very limited. As far we can determine, our methodology is novel, and any independent association with adverse health outcome, whilst intuitively expected, has never been systematically determined.
Methods

Population studied

All patients, including children and young people, registered with a Wolverhampton General Practitioner (GP) and those who were residents known to have had any hospital contact.

Data

Individual data were anonymised. To identify the link between broadband provision, demographic and health outcomes, a single dataset was developed by merging the Point Topic Broadband Digital Exclusion and the Wolverhampton Integrated Health Care databases using post codes to which local authority area codes for the Lower Layer Super Output Areas (LLSOAs) are mapped. LLSOAs describes a geographical breakdown of the UK, with an average population of 1500 people or 650 households, and are widely used to improve the reporting of small populations in NHS data modelling.

The Wolverhampton Integrated Healthcare database

The Wolverhampton integrated healthcare database is a database developed to integrate data from primary care, community and hospital clinical and pathology systems, for the residents of Wolverhampton, and from immediately adjacent districts, for all people residing in Wolverhampton or registered to Wolverhampton practices, and from immediately adjacent districts with emergency admission to New Cross Hospital (NXH), using methodology we have previously described.

The demographic and health care variables extracted were age, gender, ethnicity, the Index of Multiple Deprivation (IMD) ranked score as determined from postcodes, long term condition comorbidities and non-elective health activity. Ethnicity data from all sources were reviewed, only unambiguous data were accepted, then recoded into White, South-Asian, Black, Mixed Ethnicity, Chinese or Unknown. The sum of comorbidities in any individual was determined by the total number of comorbidities which matched the long-term conditions found in the population which are utilised in defining comorbidity in the local health economy commonest comorbidities. These are: asthma, atrial fibrillation, cancer, cardiovascular diseases, chronic kidney disease, chronic obstructive pulmonary disease, dementia, depression and other mental health
disorder, diabetes, epilepsy, haemoglobinopathy, heart failure, hypertension, learning difficulties, osteoarthritis and rheumatoid arthritis. The sum of non-elective health activity events was determined by the number of emergency department attendances not leading to admission plus the number of non-elective admissions over the preceding 12 months to our local hospital.

**The Point Topic database and the identification of appropriate measure of broadband provision**

The Point Topic® system develops a global Broadband Digital Deprivation Index from a variety of variables including the Broadband Infrastructure Index (BII) but also the IMD summary score and other ranked measures derived from the IMD such as age, income and housing. The index is a ranked score for all English LLSOAs. The BII subcomponent was the variable selected as a measure of broadband infrastructure provision. The BII is an index measure representative of broadband provision including the type, speed and bandwidth quality and number of suppliers for broadband in those defined localities\(^{17, 20}\). The BII assessment used is for the year 2020. Thus, within the merged dataset, there were a number of variables that might have shared common subcomponents and particular care was taken to ensure the independence of BII as a variable (presented in Results). We emphasise the BII to be a measure of provision, not access or uptake and, just as standard practice with the use of IMD, it defines the characteristic of a geographical unit, the LLSOA, not the individuals who may live in it.

**Statistical method**

All data were analysed on IBM SPPS version 26. When comparing independent groups, the Students t-test and the Chi-square test were used for the difference between means and proportions respectively. Principle components analysis (with rotation) was utilised when considering the inter-dependence of independent variables especially in relationship to the BII.

The BII, an ordinal variable, was not normally distributed and clearly dichotomised into two groups above and below a rank value of 10,000 (Figure 1). The purpose of this analysis was to define associations with the poorer provisioned broadband group with a BII of <10,000. Thus, binary logistic regression analysis was used to examine the relationship between BII as the dependent (dichotomised) variable and various independent variables. The continuous independent variables of IMD and age, were categorised on the basis of quintiles.
and thus utilised as categorical variables. Comorbidities was categorised as those with a total sum comorbidity score less than three and a total sum score of greater than or equal to three. The independent variables were in two broad groupings 1) the demographic variables of age, gender, ethnicity, and 2) variables of deprivation; the two health outcome variables of long-term condition comorbidity and of emergency non-elective urgent care. There is a known association and thus a potential confounding effect between these two independent variable groupings and this was negated in our analysis by taking a stepwise approach, entering each block separately as previously described. The regression analysis was undertaken on 250,609 residents as 19,176 were exclude by missing or unclassifiable ethnicity coding. Results are presented as the mean ± SD or as percentages. Statistical significance was taken at p<0.05.

Ethical Approval

Not applicable.

As the data were accrued and analysed to determine the service imperative to be aware of digital exclusion as a potential factor in delivering equitable social and health care, and as no selection, or randomisation or intervention occurred, research ethical approval was not deemed necessary, and this was confirmed within local governance processes.

Patient and Public Involvement

None as not applicable to this type of study.
Results

Population and Localities

The final cohort were the 269,785 alive at the point of data acquisition in the year 2020 and included 13,945 residents registered to a non-Wolverhampton GP (Table 1). The datasets were linked on locality codes. There are 158 LLSOAs in Wolverhampton with a surface area of (mean ±SD) (range) 0.44 ±0.28 (0.13 – 2.20) km², with 691 ±139 (466 – 1388) and 1708±435 (857 – 3476) households and individuals per locality respectively.

We again emphasize that IMD and BII data are derived from rankings relating to the locality codes and not from direct assessment of individuals and we do not have individual socio-economic data.

The Broadband Infrastructure Index (BII)

The wider Broadband Digital Deprivation Index is derived from many variables including those related to deprivation. Thus, the correlation between IMD (as IMD rank score) and the global Broadband Digital Deprivation Index was tight and highly significant (r = 0.902, p<0.001). In a principal component analysis, including all subsidiary variables, four components emerged (Eigen value of >=1), the BII sub-component of the Broadband Digital Deprivation Index was independent of all other IMD indices apart from IMD housing rankings, which equally spanned the deprivation and BII components, which has face validity. Accordingly, the linear correlation between BII and IMD ranking was extremely weak (r = -0.16, p<0.001) with an r² of only 3%, confirming with high confidence that the two measures were independent. The BII component was therefore identified as an appropriate measure of broadband provision for linkage to health data at the LLSOA local authority area code levels. We emphasise BII not to be a continuous variable since the group was clearly divisible into two cohorts of those with a distinctly low BII versus others at a threshold rank score of 10,000 (Figure 1).

The association between broadband provision, demographic and health outcomes

The demographic and health data of the whole population and of the BII groupings are shown in Table 1. Regarding demography, the lower scoring BII group (lesser provision) were older and also had marginally more females. Of particular note, for the lesser provisioned BII category, both the IMD parameters indicated
this group to be less, not more, deprived and they had a higher, not lower, white ethnicity prevalence. The crude relationship of the BII as a dichotomised group to quintiles of the IMD score ($\chi^2=7559.4$, p<0.001), quintiles of age ($\chi^2=156.7$, p<0.001) and with ethnicity ($\chi^2=1521.5$, p<0.01) is depicted in Figure 2. For parameters of health outcomes (Table 1), the lower scoring BII group was significantly more comorbid as measured by the total number of comorbidities or by the proportion with any comorbidity and they had a very small but significant excess of non-elective urgent health activity over the preceding 12 months. Since BII is not a continuous variable, these factors were also entered into binary logistic regression of the dichotomised BII status as the dependent variable with the purpose of highlighting the magnitude, direction and significance of any association with the lower BII group of the defined independent variables which were entered in two stages: firstly, the demographic factors of quintiles of age, gender (male vs female), ethnicity (white vs other), and quintiles of the IMD score; secondly, comorbidity (any versus none) and urgent events (any versus none). Table 2 shows the outputs of this model which was highly significant ($\chi^2 = 7,914.8$, p<0.001), albeit the independent factors explained very little of the variance ($r^2 = 0.042$). At step 1, the model adjusted Odds Ratios confirmed the crude associations with IMD (p<0.001) and ethnicity (p<0.001) but exposed a different (model adjusted) pattern for age (p<0.001) with the younger quintiles to be in the low scoring BII group whilst the association of gender was not significant. At step 2, the relationship of health indices of any comorbidity (p<0.001) and any urgent care event (p<0.001) to low BII were both highly significant, noting these two latter variables to have a degree of correlation ($\chi^2 = 3719.4$, Eta 0.12, p<0.001).
Table 1: Demographic and health data of the whole resident population, subdivided by categories of the Broadband Infrastructure Index (BII).

Results are the mean ± SD with (range), ns = non-significant and the p-value is for the significance of the difference between the two BII categories.

<table>
<thead>
<tr>
<th></th>
<th>Whole population n= 269,785</th>
<th>BII &lt; 10,000 (worse) n=125,007 (46%)</th>
<th>BII&gt;10,000 n=144,778 (54%)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BII rank</td>
<td>11,657±6,926</td>
<td>4,971±348 (3,451 – 5,095)</td>
<td>17,430±4168 (11,328 – 27,276)</td>
<td>t=1,053.6, p&lt;0.001</td>
</tr>
<tr>
<td>Age (years)</td>
<td>38.6±23.2 (0 – 106)</td>
<td>39.1±23.3</td>
<td>38.3±23.0</td>
<td>t=8.816, p&lt;0.001</td>
</tr>
<tr>
<td>Gender (male %)</td>
<td>50.4%</td>
<td>50.1%</td>
<td>50.6%</td>
<td>χ² = 4.45, p&lt;0.05</td>
</tr>
<tr>
<td>Ethnicity White %</td>
<td>57.1%</td>
<td>65.6%</td>
<td>57.9%</td>
<td>χ² = 1521.5, p&lt;0.001</td>
</tr>
<tr>
<td>IMD Score</td>
<td>34.7±15.9 (5.4 – 71.8)</td>
<td>32.6±15.6</td>
<td>36.4±15.8</td>
<td>t=63.549, p&lt;0.001</td>
</tr>
<tr>
<td>IMD Rank</td>
<td>8878±7872</td>
<td>9698±7739</td>
<td>8169±7916</td>
<td>t=50.586, p&lt;0.001</td>
</tr>
<tr>
<td>Number of comorbidities</td>
<td>0.61±1.06 (0 – 11)</td>
<td>0.63±1.07</td>
<td>0.59±1.05</td>
<td>t=11.069, p&lt;0.001</td>
</tr>
<tr>
<td>Any comorbidity</td>
<td>35%</td>
<td>36.5%</td>
<td>34%</td>
<td>χ² = 173.54, p&lt;0.001</td>
</tr>
<tr>
<td>Number of non-elective contacts in 12 months</td>
<td>0.24±0.84 (0 – 65)</td>
<td>0.242±0.85</td>
<td>0.235±0.83</td>
<td>t=2.042, p&lt;0.05</td>
</tr>
<tr>
<td>Any non-elective contact in 12 months</td>
<td>14.4 %</td>
<td>14.7%</td>
<td>14.4%</td>
<td>χ² = 16.01, p&lt;0.001</td>
</tr>
</tbody>
</table>
Table 2: Binary logistic regression identifying factors associated with the poor or least provisioned Broadband Infrastructure Index (BII) group.

The analysis was undertaken on 250,609 subjects with 19,176 missing cases due to a lack of ethnicity data. The Odds ratios (OR) are for the poor or least provisioned BII group. IMD is the Index of Multiple Deprivation. An * represents the reference group for each categorical variable from which all other groups in that variable were compared.

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR (95% CI)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMD Q1 (11.6±3.2) (least deprived)</td>
<td>2.14 (2.09-2.19)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>IMD Q2 (24.3±4.0)</td>
<td>2.37 (2.31-2.43)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>IMD Q3 (36.8±2.5)</td>
<td>1.38 (1.35-1.42)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>IMD Q4 (44.1±2.0)</td>
<td>1.27 (1.24-1.31)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>IMD Q5 (56.1±5.2) (most deprived)</td>
<td>1.0</td>
<td>-</td>
</tr>
<tr>
<td>Age Q1 (7.7±4.5 years) (youngest)</td>
<td>1.20 (1.16-1.23)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age Q2 (24.0±4.7 years)</td>
<td>1.08 (1.05-1.11)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age Q3 (38.3±4.0 years)</td>
<td>1.08 (1.05-1.11)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age Q4 (52.8±4.3 years)</td>
<td>1.06 (1.03-1.09)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age Q5 (72.8±8.7 years) (oldest)</td>
<td>1.0</td>
<td>-</td>
</tr>
<tr>
<td>White Ethnicity (61.5%)</td>
<td>1.33 (1.31-1.36)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Other ethnic groups *</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td><strong>Step 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any comorbidities present (35.2%)</td>
<td>1.07 (1.05-1.09)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No comorbidities*</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Any emergency episode (14.4%)</td>
<td>1.05 (1.03-1.08)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>No emergency episodes*</td>
<td>1.0</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1 Demonstrating the dichotomised distribution of the Broadband Infrastructure Index (BII) and its suitability for binary logistic regression analysis.

Figure 2 The crude relationships of IMD, age and Ethnicity to the category of the Broadband Infrastructure Index (BII).
Discussion

On the contention that physical availability of broadband infrastructure should not be overlooked as a precursor to health care access\textsuperscript{22}, we aimed to identify the nature of association between broadband provision, population demographics and health outcomes in a single health economy, ours being deprived and multi-ethnic by UK standards. To our knowledge, this is the first study to identify a candidate measure of broadband provision (BII) and link it to population-level health data (urgent care and comorbidity) and indices of deprivation (IMD). In some sense, we wished to consider the epidemiology of broadband provision. We find that associations exist between BII determined broadband provision, demographic variables and health outcomes.

The derivation of the BII is described elsewhere\textsuperscript{17, 20}. Previous studies, to the best of our knowledge, have not identified or utilised any adequate measure of broadband infrastructure\textsuperscript{13, 23-25} and we are the first to demonstrate that it is essentially an independent factor so long as it is disaggregated from association with deprivation data when used in this context. Had we used the broader more global Broadband Digital Deprivation Index as the measure, these associations would not have been seen, confounded by strong collinearity with IMD. BII in its own right can be utilised to demonstrate variation by geographical locality. Our own data showing localities that stood out as less well provisioned emphasises the point that any such avoidable digital divide is an addressable infrastructure issue. How such variation arose is unknown to us, perhaps by historical chance events, or other practical infrastructure constraints, or because broadband providers are driven by commercial and fiscal but not social imperatives\textsuperscript{24}. Nevertheless, geographical variation of provision can only be addressed if measured.

We demonstrate in the crude data and in regression models that BII was associated with variation with demography and with measurable health outcomes. In binary logistic regression, contrary to expectation, the lower provisioned BII group were less deprived, had a lower non-white ethnicity make-up, and the adjusted association with age showed an impact in younger age groups. The association of poorer broadband provision and the health outcomes of both comorbidity and non-elective urgent care has not been previously described at a population level. There arise a number of considerations: whilst the associations were highly statistically significant the size of the effect was very small with the independent variables describing < 5% of the variance of BII. We are thus simply stating that they can exist. We believe our local findings are unlikely to be definitive and will vary from region to region, the point being that, unless measured, accountable organisations cannot
be assured of equitable provision as a platform for equitable access other than perhaps by geography. Regarding geography, published comparisons are generally made between large population blocks at regional or national level\textsuperscript{16, 20, 26}, whereas we have considered broadband provision within a single local healthcare economy drilling down to the LLSOAs\textsuperscript{20} and determined that focus at that level ought to be a consideration for local municipal and health authorities.

The distinction between provision and uptake also requires emphasis, perhaps crudely summarised as the distinction between public provision and individual access. The latest UK government and other independent analysis shows rapidly increasing internet usage\textsuperscript{16, 20, 26} with less than 10\% of the population being non or infrequent users, although poor usage remains associated with lower income, lesser educational attainment, increasing age, to a much lesser extent ethnicity and gender as well as geographical factors including the rural urban divide. At an individual level, the skill set required for internet usage and the nature of that use is increasingly understood\textsuperscript{26}, noting that internet use for health remains relatively low. Regarding the skill set required, 8\% of the UK population have no basic level skills and in 21\% they are limited\textsuperscript{27}. Generally, there is evidence that broadband and internet usage are associated with a variety of benefits amongst those who are digitally included\textsuperscript{27, 28}. The concept of the digital divide and digital exclusion arises out of a concern for the capability among individuals and communities to benefit from digital innovations. In health care, specific measures have been developed to assess digital literacy including e-Heals\textsuperscript{29} and applied in differing digital health interventions. From that has arisen an understanding of the barriers to engagement with such digital health interventions\textsuperscript{30}, how and whether health behaviour can be modified\textsuperscript{31} and, importantly, that digital health interventions, of themselves, can lead to inequality\textsuperscript{32}.

Publicly accessible digital facilities and personal smartphones may be solutions to inequitable home broadband provision, although they have limitations\textsuperscript{20}. Broadband access within an individual’s home, rather than a community space, is thought important for increasing digital engagement and associated with higher satisfaction, a wider freedom of use and more complex activities including those related to healthcare\textsuperscript{33-35}. However, bridging the gap from infrastructure to broadband access and internet use, end point access, previously often described as ‘last mile access’, is likely determined by uptake. As broadband infrastructure provision increasingly becomes a politically recognised social determinant of health and other outcomes\textsuperscript{4, 36, 37, 38, 39}, with reports describing broadband infrastructure as a “super social determinant of health”\textsuperscript{23, 40}, it perhaps will come to be recognised as a basic social right, especially if there is an overlap between broadband provision and social inequality\textsuperscript{1, 12, 41}. If broadband provision is an investment in the health and wealth of a population,
then digital poverty might require greater recognition, for which a direct analogy exists for the provision of energy utilities, fuel poverty and fuel payments\cite{42}.

**Caveats, Strengths and Weaknesses**

We recognise that the BII is a geographical parameter and does not tell us about individual digital exclusion, since we do not have such assessments, just as the use of IMD is taken in lieu of individual detailed socio-economic data.

Our findings are pertinent to our local health economy and may well differ from similarly sized unitary health and social care geopolitical areas.

This study is a cross-sectional time point view in rapidly changing digital environment.

As a further caveat we emphasise that only a prospective follow up study can determine whether broadband provision is able to modify health determinants and outcomes, which surely must be the intent of all digital health initiatives.

With those caveats and weaknesses, we believe this to be the first description of the association between measurable broadband provision and health data at a population level and we are the first to demonstrate a statistically significant link with certain crucial health outcomes - comorbidity and urgent care.

Our study methodology is generalisable and based on a high quality and complete dataset curated for the local population. This enables assessment at the local level, taking into account local variations and moving away from large scale inter-regional or national level comparisons.

**Conclusion**
The provision of broadband is but one component of the digital infrastructure capacity and capability required for an individual to access and reap the potential benefits of the digitalisation health and social care. We describe a methodology to measure such provision and demonstrate that its variation can be determined at a very local level within unitary and health and social care economies. Such differences may be associated with an unintended variation in a populations’ demography and with important health imperatives. Such associations cannot be assumed but are best measured.

Implications for clinicians and policymakers

Local authorities may wish to consider and extend our described methodology to ensure equitable provision of broadband infrastructure.

Individual care providing services and their staff must be aware of the potential constraints faced by their clients and patients as digital interventions are developed as well as their own limitations\(^{13,43}\).

Future research

As is the intent of the WODEN group, hopefully, our work will promote future research in this field to address the many aspects needed to ensure equity in health and care delivery, and the demonstrable avoidance of imparting disequity, as we so rapidly move into the digital paradigm.

Author contributions

Accountable senior author BMS; VK; Manuscript writing: FP, AFN, VK, KW, EP, HC, JB, OJ, AN, ADP, BMS; Preparation for submission: FP, AFN, AN, ADP, BMS; Database quality, data integration and data quality and integration: BMS, VK; Statistical advice: AN, BMS, All authors contributed intellectual content during the drafting and revision of the work and approved the final version.

Transparency declaration
BMS affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

**Funding**

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interest statement**

All authors have completed the Unified Competing Interest form (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years, no other relationships or activities that could appear to have influenced the submitted work.

**Data Sharing**

Anonymised data will be shared on reasonable request to the corresponding author.

**Ethics statement**

Not applicable.

As the data were accrued and analysed to determine the service imperative to be aware of digital exclusion as a potential factor in delivering equitable social and health care, and as no selection, or randomisation or intervention occurred, research ethical approval was not deemed necessary, and this was confirmed within local governance processes.
References


The crude relationships of IMD, Age and Ethnicity to the category of the Broadband Infrastructure Index (BII)
<table>
<thead>
<tr>
<th>Item No.</th>
<th>Recommendation</th>
<th>Page No.</th>
<th>Relevant text from manuscript</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1</strong></td>
<td><em>(a)</em> Indicate the study’s design with a commonly used term in the title or the abstract</td>
<td>1</td>
<td>An observational cross-sectional cohorted whole population study of the association of poor broadband provision with demographic and health outcomes as determined in a single deprived multi-ethnic health economy: The Wolverhampton Digital Enablement (WODEN) Programme</td>
</tr>
<tr>
<td></td>
<td><em>(b)</em> Provide in the abstract an informative and balanced summary of what was done and what was found</td>
<td>3</td>
<td>See abstract</td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>Explain the scientific background and rationale for the investigation being reported</td>
<td>5</td>
<td>Published research that systematically analyses the association between impaired digital access with health outcomes is very limited. As far we can determine, our methodology is novel, and any independent association with adverse health outcome, whilst intuitively expected, has never been systematically determined.</td>
</tr>
<tr>
<td><strong>3</strong></td>
<td>State specific objectives, including any prespecified hypotheses</td>
<td>5</td>
<td>The aim of this study was to focus on broadband provision and to consider, at population scale, its association with</td>
</tr>
</tbody>
</table>
## Methods

<table>
<thead>
<tr>
<th>Study design</th>
<th>4</th>
<th>Present key elements of study design early in the paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Setting</td>
<td>5</td>
<td>Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection</td>
</tr>
</tbody>
</table>
| Participants | 6 | (a) **Cohort study**—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up  
Case-control study—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  
Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants |
|  | | (b) **Cohort study**—For matched studies, give matching criteria and number of exposed and unexposed  
Case-control study—For matched studies, give matching criteria and the number of controls per case |
| Variables | 7 | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable |
| 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 |
| Bias | 9 | Describe any efforts to address potential sources of bias |

There were a number of variables that might have shared common subcomponents and particular care was taken to ensure the independence of BII as a variable (presented in Results). We emphasise the BII to be a measure of provision.
not access or uptake and, just as standard practice with the use of IMD, it defines the characteristic of a geographical unit, the LLSOA, not the individuals who may live in it.

There is a known association and thus a potential confounding effect between these two independent variable groupings and this was negated in our analysis by taking a stepwise approach, entering each block separately as previously described.

<table>
<thead>
<tr>
<th>Study size</th>
<th>10</th>
<th>Explain how the study size was arrived at</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6,9</td>
<td>All patients registered with a Wolverhampton General Practitioner (GP) and those who were residents known to have had any hospital contact.</td>
</tr>
</tbody>
</table>

The final cohort were the 269,785 alive at the point of data acquisition in the year 2020 and included 13,945 residents registered to a non-Wolverhampton GP (Table 1).
<table>
<thead>
<tr>
<th>Quantitative variables</th>
<th>11</th>
<th>Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistical methods</td>
<td>12</td>
<td>(a) Describe all statistical methods, including those used to control for confounding</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(b) Describe any methods used to examine subgroups and interactions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(c) Explain how missing data were addressed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(d) Cohort study—If applicable, explain how loss to follow-up was addressed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Case-control study—If applicable, explain how matching of cases and controls was addressed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cross-sectional study—If applicable, describe analytical methods taking account of sampling strategy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(e) Describe any sensitivity analyses</td>
</tr>
<tr>
<td>Results</td>
<td></td>
<td>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</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(b) Give reasons for non-participation at each stage</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(c) Consider use of a flow diagram</td>
</tr>
<tr>
<td>Descriptive data</td>
<td>14*</td>
<td>(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(b) Indicate number of participants with missing data for each variable of interest</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(c) Cohort study—Summarise follow-up time (eg, average and total amount)</td>
</tr>
<tr>
<td>Outcome data</td>
<td>15*</td>
<td>Cohort study—Report numbers of outcome events or summary measures over time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Case-control study—Report numbers in each exposure category, or summary measures of exposure</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cross-sectional study—Report numbers of outcome events or summary measures</td>
</tr>
<tr>
<td>Main results</td>
<td>16</td>
<td>(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</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(b) Report category boundaries when continuous variables were categorized</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period</td>
</tr>
</tbody>
</table>

The regression analysis was undertaken on 250,609 residents as 19,176 were exclude by missing or unclassifiable ethnicity coding.
Continued on next page
<table>
<thead>
<tr>
<th>Other analyses</th>
<th>17</th>
<th>Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses</th>
<th>6 to 13</th>
</tr>
</thead>
</table>

**Discussion**

<table>
<thead>
<tr>
<th>Key results</th>
<th>18</th>
<th>Summarise key results with reference to study objectives</th>
<th>14 to 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limitations</td>
<td>19</td>
<td>Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias</td>
<td>15 to 16</td>
</tr>
<tr>
<td>Interpretation</td>
<td>20</td>
<td>Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence</td>
<td>14 to 16</td>
</tr>
<tr>
<td>Generalisability</td>
<td>21</td>
<td>Discuss the generalisability (external validity) of the study results</td>
<td>14 to 16</td>
</tr>
</tbody>
</table>

**Other information**

<table>
<thead>
<tr>
<th>Funding</th>
<th>22</th>
<th>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</th>
<th>17</th>
</tr>
</thead>
</table>

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.*

**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 www.strobe-statement.org.

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.