Green–blue space exposure changes and impact on individual-level well-being and mental health: a population-wide dynamic longitudinal panel study with linked survey data

Rebecca S Geary, Daniel A Thompson, Joanne K Garrett, Amy Mizen, Francis M Rowney, Jiao Song, Mathew P White, Rebecca Lovell, Alan Watkins, Ronan A Lyons, Susan Williams, Gareth Stratton, Ashley Akbari, Sarah C Parker, Mark J Nieuwenhuijsen, James White, Benedict W Wheeler, Richard Fry, Dialechti Tsimpida and Sarah E Rodgers

Green–blue space exposure changes and impact on individual-level well-being and mental health: a population-wide dynamic longitudinal panel study with linked survey data

Rebecca S Geary,[1] Daniel A Thompson,[2]
Joanne K Garrett,[3] Amy Mizen,[2] Francis M Rowney,[3]
Jiao Song,[2] Mathew P White,[3] Rebecca Lovell,[3]
Alan Watkins,[2] Ronan A Lyons,[2] Susan Williams,[4]
Gareth Stratton,[5] Ashley Akbari,[2] Sarah C Parker,[2]
Mark J Nieuwenhuijsen,[6] James White,[7]
Benedict W Wheeler,[3] Richard Fry,[2]
Dialechti Tsimpida[1] and Sarah E Rodgers[1*]

[1]Department of Public Health, Policy and Systems, University of Liverpool, Liverpool, UK
[2]Department of Health Data Science, Swansea University, Swansea, UK
[3]European Centre for Environment and Human Health, University of Exeter Medical School, University of Exeter, Truro, UK
[4]Natural Resources Wales, Bangor, UK
[5]College of Engineering, Swansea University, Swansea, UK
[6]ISGlobal, Barcelona, Spain
[7]Centre for Trials Research, School of Medicine, Cardiff University, Cardiff, UK

[*]Corresponding author

Disclosure of interests

Public Health Research

Criteria for inclusion in the *Public Health Research* journal

Reports are published in *Public Health Research* (PHR) if (1) they have resulted from work for the PHR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Public Health Research* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

PHR programme

The Public Health Research (PHR) programme, part of the National Institute for Health and Care Research (NIHR), is the leading UK funder of public health research, evaluating public health interventions, providing new knowledge on the benefits, costs, acceptability and wider impacts of non-NHS interventions intended to improve the health of the public and reduce inequalities in health. The scope of the programme is multi-disciplinary and broad, covering a range of interventions that improve public health.

For more information about the PHR programme please visit the website: https://www.nihr.ac.uk/explore-nihr/funding-programmes/public-health-research.htm

This report

Abstract

Green–blue space exposure changes and impact on individual-level well-being and mental health: a population-wide dynamic longitudinal panel study with linked survey data

Rebecca S Geary,[1] Daniel A Thompson,[2] Joanne K Garrett,[3] Amy Mizen,[2] Francis M Rowney,[3] Jiao Song,[2] Mathew P White,[3] Rebecca Lovell,[3] Alan Watkins,[2] Ronan A Lyons,[2] Susan Williams,[4] Gareth Stratton,[5] Ashley Akbari,[2] Sarah C Parker,[2] Mark J Nieuwenhuijsen,[6] James White,[7] Benedict W Wheeler,[3] Richard Fry,[2] Dialechti Tsimpida[1] and Sarah E Rodgers[1*]

[1]Department of Public Health, Policy and Systems, University of Liverpool, Liverpool, UK
[2]Department of Health Data Science, Swansea University, Swansea, UK
[3]European Centre for Environment and Human Health, University of Exeter Medical School, University of Exeter, Truro, UK
[4]Natural Resources Wales, Bangor, UK
[5]College of Engineering, Swansea University, Swansea, UK
[6]ISGlobal, Barcelona, Spain
[7]Centre for Trials Research, School of Medicine, Cardiff University, Cardiff, UK

[*]Corresponding author Sarah.Rodgers@liverpool.ac.uk

Background: Cross-sectional evidence suggests that living near green and blue spaces benefits mental health; longitudinal evidence is limited.

Objectives: To quantify the impact of changes in green and blue spaces on common mental health disorders, well-being and health service use.

Design: A retrospective, dynamic longitudinal panel study.

Setting: Wales, UK.

Participants: An e-cohort comprising 99,682,902 observations of 2,801,483 adults (≥ 16 years) registered with a general practice in Wales (2008–2019). A 5312-strong 'National Survey for Wales (NSW) subgroup' was surveyed on well-being and visits to green and blue spaces.

Main outcome measures: Common mental health disorders, general practice records; subjective well-being, Warwick–Edinburgh Mental Well-being Scale.

Data sources: Common mental health disorder and use of general practice services were extracted quarterly from the Welsh Longitudinal General Practice Dataset. Annual ambient greenness exposure, enhanced vegetation index and access to green and blue spaces (2018) from planning and satellite data. Data were linked within the Secure Anonymised Information Linkage Databank.

Methods: Multilevel regression models examined associations between exposure to green and blue spaces and common mental health disorders and use of general practice. For the National Survey for Wales subgroup, generalised linear models examined associations between exposure to green and blue spaces and subjective well-being and common mental health disorders.

Results and conclusions: Our longitudinal analyses found no evidence that changes in green and blue spaces through time impacted on common mental health disorders. However, time-aggregated exposure to green and blue spaces contrasting differences between people were associated with subsequent common mental health disorders. Similarly, our cross-sectional findings add to growing evidence that residential green and blue spaces and visits are associated with well-being benefits:

- Greater ambient greenness (+ 1 enhanced vegetation index) was associated with lower likelihood of subsequently seeking care for a common mental health disorder [adjusted odds ratio (AOR) 0.80, 95% confidence interval, (CI) 0.80 to 0.81] and with well-being with a U-shaped relationship [Warwick–Edinburgh Mental Well-being Scale; enhanced vegetation index beta (adjusted) −10.15, 95% CI −17.13 to −3.17; EVI^2 beta (quadratic term; adj.) 12.49, 95% CI 3.02 to 21.97].
- Those who used green and blue spaces for leisure reported better well-being, with diminishing extra benefit with increasing time (Warwick–Edinburgh Mental Well-being Scale: time outdoors (hours) beta 0.88, 95% CI 0.53 to 1.24, time outdoors^2 beta −0.06, 95% CI −0.11 to −0.01) and had 4% lower odds of seeking help for common mental health disorders (AOR 0.96, 95% CI 0.93 to 0.99).
- Those in urban areas benefited most from greater access to green and blue spaces (AOR 0.89, 95% CI 0.89 to 0.89).
- Those in material deprivation benefited most from leisure time outdoors (until approximately four hours per week; Warwick–Edinburgh Mental Well-being Scale: time outdoors × in material deprivation: 1.41, 95% CI 0.39 to 2.43; time outdoors^2 × in material deprivation −0.18, 95% CI −0.33 to −0.04) although well-being remained generally lower.

Limitations: Longitudinal analyses were restricted by high baseline levels and limited temporal variation in ambient greenness in Wales. Changes in access to green and blue spaces could not be captured annually due to technical issues with national-level planning datasets.

Future work: Further analyses could investigate mental health impacts in population subgroups potentially most sensitive to local changes in access to specific types of green and blue spaces. Deriving green and blue spaces changes from planning data is needed to overcome temporal uncertainties.

Contents

CONTENTS

CONTENTS

List of tables

List of figures

List of supplementary material

Report Supplementary Material 1　Read codes used in algorithm to extract specific GPE relating to common mental health disorders

Report Supplementary Material 2　Supplementary results tables for Chapter 7

Report Supplementary Material 3　Guidance for Reporting Involvement of Patients and the Public 2 (short form)

Supplementary material can be found on the NIHR Journals Library report page (https://doi.org/10.3310/LQPT9410).

Supplementary material has been provided by the authors to support the report and any files provided at submission will have been seen by peer reviewers, but not extensively reviewed. Any supplementary material provided at a later stage in the process may not have been peer reviewed.

List of abbreviations

ALF	anonymised linking field	NIR	near-infrared
AIRR	adjusted incidence rate ratio	NRW	Natural Resources Wales
AOR	adjusted odds ratio	NSW	National Survey for Wales
CI	confidence interval	ONS	Office for National Statistics
CMD	common mental health disorders	OD	origin-destination
EHR	electronic health record	OR	odds ratio
EVI	enhanced vegetation index	OS	Ordnance Survey
GIS	geographic information system	OSMM	Ordnance Survey MasterMap®
GAM	generalised additive model	PHR	Public Health Research
GBS	green and blue spaces	RALF	Residential Anonymous Linking Field
GLM	generalised linear model	SAIL	Secure Anonymised Information Linkage
GP	general practitioner	SD	standard deviation
GPE	general practice events	TAN	Technical Advice Note
HEPE	health and environment public engagement	WDSD	Welsh Demographic Service Dataset
IRR	incidence rate ratio	WEMWBS	Warwick–Edinburgh Mental Well-being Scale
LSOA	lower-layer super output area	WHO	World Health Organization
MHI-5	Mental Health Inventory 5	WIMD	Welsh Index of Multiple Deprivation
NDVI	normalised difference vegetation index	WLGP	Welsh Longitudinal General Practice
NIHR	National Institute for Health and Care Research		

Plain language summary

We investigated whether people who live near or visit green (parks, woodlands) and blue (riversides, beaches) spaces have fewer common mental health disorders (anxiety or depression), and better well-being. We considered whether changes in the amount of green and blue space around the home affected people's mental health.

We assessed the availability of local green and blue spaces. Annual exposure and access to local green and blue spaces were extracted from planning and satellite data. We linked these data to anonymised health records of 2,801,483 adults registered with a general practice from 2008 to 2019, and to survey answers about leisure visits to natural environments and well-being. We found:

- people who lived in greener and bluer areas were less likely to seek help for a common mental health disorder than those in less green or blue areas, with those living in the most deprived areas benefiting the most
- people who used green and blue spaces for leisure, especially those with the greatest levels of deprivation, had better well-being and were less likely to seek help for common mental health disorders
- no evidence that changing amounts of green and blue space affected how likely people were to seek help for common mental health disorders; this may be because we found mostly small changes in green and blue space, and we may not have allowed enough time between moving home and recording mental health.

We found evidence for relationships between green and blue space and mental health. However, some analyses were restricted due to lack of data on changes in green and blue spaces. An important finding was that people in deprived communities appear to benefit the most. Provision of green and blue spaces could be a strategy to improve the mental health of people living in disadvantaged areas.

Scientific summary

Background

Cross-sectional evidence suggests that living close to and spending time in and around green and blue spaces (GBS, e.g. parks, gardens, ponds, lakes, rivers) is associated with higher levels of well-being and fewer mental health problems. Some of these benefits could result from living near GBS; others (e.g. physical activity) are attributed to using GBS. There are few longitudinal studies examining whether changes in GBS affect subsequent mental health. Existing studies have largely had small cohorts, assessed change over a short time, did not investigate socioeconomic inequalities in effects and neglected to consider that some health benefits depend on using GBS rather than simple exposure.

Objectives

We examined the research questions:

1. Is a greater residential exposure and access to GBS associated with the likelihood of seeking help for a common mental health disorder (CMD) in general practice?
2. Is the association between exposure, or access to, GBS and seeking help for a CMD modified by socioeconomic disadvantage and moving home?
3. Is a greater residential exposure associated with the likelihood of using general practice services?
4. Is greater residential GBS exposure and access to GBS associated with higher subjective well-being or a lower likelihood of seeking help for CMD, and does intentional use of GBS for leisure mediate these associations?
5. Are associations between residential GBS exposure and access to GBS and use and subjective well-being mediated by walking-specific physical activity and modified by socioeconomic disadvantage?

Methods

We used a dynamic longitudinal panel study design to retrospectively examine natural and non-random variation in GBS exposure during an 11-year period, with individual linkage to survey responses.

Data sources

We constructed an electronic cohort of the population of Wales using data from the Welsh Demographic Service Dataset. Household-level ambient greenness was determined from satellite imagery, access to GBS from planning records, CMD and service use from Welsh Longitudinal General Practice Dataset (WLGPD) and for a subgroup, well-being and behaviour responses to the National Survey for Wales (NSW, 2016/17 and 2018/19). Data are held in the Secure Anonymised Information Linkage (SAIL) databank.

Inclusion criteria

The cohort included everyone aged 16 years and over registered with a general practitioner (GP) providing patient records to the SAIL databank between January 2008 and October 2019 ('the study period'). Those not registered with an appropriate GP, not having a Welsh residential address between January 2008 and October 2019, and not having sex/week of birth recorded were excluded from the study sample.

Exposures

- Our primary environmental exposure was ambient greenness, measured yearly (2008–19) by mean enhanced vegetation index (EVI) averaged over a 300 m circular buffer centred on each residence.
- Our secondary environmental exposure was the potential for an individual to access GBS along paths/roads within 1600 m of each home, modelled for 2018. We captured changes only when individuals moved home. For research questions 1–3, GBS access was quantified by the number of potentially accessible GBS. For research questions 4–5, access was quantified as the proximity to the nearest GBS.

Outcome measures

- For research questions 1–3, our primary outcome was seeking help for CMD in general practice, identified from practice records using a validated algorithm. Our secondary outcome was the quarterly count of general practice events.
- For research questions 4 and 5, our primary outcome was subjective well-being, measured by the Warwick–Edinburgh Mental Well-being Scale (WEMWBS) in the NSW. Secondary outcomes were life satisfaction (NSW) and seeking help for CMD in general practice. Self-reported time outdoors on recreational visits to GBS in Wales and walking-specific physical activity were identified (NSW).

Statistical analysis

Research questions 1–3 were addressed using complementary approaches:

- We first evaluated associations between cumulative EVI exposure and seeking help for CMD in general practice using multivariate logistic regression. Follow-up stopped when help was first sought for CMD or at the end of the study period. This time aggregated design allowed us to consider environmental exposure prior to seeking help for CMD in general practice and included differences between people. We conducted stratified analyses according to area-level deprivation (Welsh Index of Multiple Deprivation, WIMD) quintiles, home moves (none/one/more than one), history of seeking care for CMD in general practice (before January 2008) and urban-rural home location.
- Next, we assessed associations between longitudinal change in GBS exposures for the same individuals, and their subsequent likelihood of seeking help for CMD (primary outcome) using multivariate logistic regression. For the association with general practice services use (secondary outcome) we used Poisson regression. We used a panel design with a multilevel model structure (quarterly observations nested within individuals). We accounted for clustering of observations over time (using random effects), and unbalanced data (number of observations varying for individuals). Change in GBS exposure was modelled as a continuous variable relative to previous time periods. For people who did not move home, we modelled EVI change categories to examine in situ exposure changes. Stratified analyses were conducted for quintiles of WIMD and number of home moves.

For research questions 4–5 we assessed whether GBS exposures were related to WEMWBS, life satisfaction and seeking help for CMD, and whether there was moderation by socioeconomic disadvantage. We examined whether the association between GBS exposures and WEMWBS, life satisfaction and seeking help for CMD, were mediated by time spent outdoors. Generalised additive models were used to identify any non-linear associations and inform subsequent generalised linear regression. When predicting WEMWBS, linear and quadratic terms for EVI were included and proximity to nearest GBS was categorised. For seeking help for CMD, EVI was categorised, and proximity was continuous. To assess moderation by socioeconomic disadvantage, we included interactions between material deprivation and the GBS exposures.

Results

The cohort included 2,801,483 individuals (99,682,902 observations). A total of 816,242 individuals (29.1%) sought help for CMD from their general practice at least once in the study period.

Research question 1: Is a greater residential exposure and access to GBS associated with the likelihood of seeking health for CMDs in general practice?

People who lived in homes surrounded by more ambient greenness (+0.1 EVI) were associated with 20% lower odds of seeking help for CMD in general practice [adjusted odds ratio (AOR) 0.80, 95% confidence interval (CI) 0.80 to 0.81]. More potentially accessible GBS (equivalent to 236 more GBS) were associated with 7% lower odds of seeking help for CMD (AOR 0.93, 95% CI 0.93 to 0.93). Every additional 360 m (0.1 unit) to the nearest GBS was associated with 5% higher odds of seeking help for CMD (AOR 1.05, 95% CI 1.04 to 1.05).

People experiencing an increase of 0.1 mean EVI was not associated with seeking help for CMD in general practice (exposure change relative to: previous quarter/year/baseline AOR 1.00, 95% CI 1.00 to 1.00). Among the 1,611,581 people who did not move home (57.5% of the cohort), increases or decreases of ≥ 0.15 EVI relative to baseline were not associated with seeking help for CMD (AOR for ≥ +0.15 EVI 1.00, 95% CI 0.99 to 1.00; AOR for ≥ −0.15 EVI 1.00, 95% CI 1.00 to 1.01). Among those who moved home at least once, an increase in the number of potentially accessible GBS was not associated with seeking help for CMD in general practice (GBS increase relative to previous quarter/year/baseline: AOR 1.00, 95% CI 1.00 to 1.00).

Research question 2: Is the association between exposure, or access to, GBS and seeking help for a CMD modified by socioeconomic disadvantage and moving home?

Living in a home with greater EVI (+0.1) and more accessible GBS (+236) was associated with lower odds of seeking help for CMD in general practice in all deprivation quintiles and whether people moved. The odds of seeking help were lower for those who did not move (AOR non-movers 0.77, 95% CI 0.77 to 0.78; AOR moved > once 0.92, 95% CI 0.91 to 0.93).

There was no evidence that EVI or GBS changes were associated with seeking help for a CMD for people living in either the most or least deprived urban areas. An increase of 0.1 mean EVI (relative to baseline) was associated with slightly lower odds of seeking help for CMD for people who moved home more than once (AOR 0.99, 95% CI 0.99 to 0.99) but not for those who moved once or had not moved.

Association between different GBS exposures and seeking help for CMD in general practice were modified by urban–rural home location and a history of seeking help for CMD. Greater EVI was associated with lower odds of seeking help for a CMD for those living in rural areas (AOR 0.59, 95% CI 0.58 to 0.59) compared with urban areas (AOR 0.85, 95% CI 0.84 to 0.85). More potentially accessible GBS were only associated with lower odds of seeking help for CMD for those living in urban areas (AOR 0.89, 95% CI 0.89 to 0.89). In rural areas more GBS access was associated with higher odds of seeking help for CMD (AOR 1.10, 95% CI 1.08 to 1.12). More ambient greenness (+0.1 EVI) was associated with lower odds of seeking help for a CMD for those who had historically sought help for CMD (before January 2008) than people who had not (AOR 0.68 vs. 0.84).

Research question 3: Is a greater residential exposure associated with the likelihood of using GP services?

There was no evidence that an increase in mean EVI was associated with the number of general practice event days/quarter (adjusted incidence rate ratio relative to baseline: 1.00, 95% CI 1.00 to 1.00).

Research question 4: Is greater residential GBS exposure and access to GBS associated with higher subjective well-being, or a lower likelihood of seeking help for CMD in general practice, and are these associations mediated by intentional use of GBS for leisure?

Among NSW respondents with outcome measures (n = 5971) EVI was significantly related to well-being as measured by WEMWBS with a U-shaped relationship (linear regression; EVI −10.15, 95% CI −17.13 to −3.17; EVI^2 12.49, 95% CI 3.02 to 21.97) and weekly time outdoors (linear relationship; EVI 43.41, 95% CI 6.85 to 79.97). Time outdoors was significantly positively related to WEMWBS [time outdoors (hours) 0.88, 95% CI 0.53 to 1.24, time outdoors (hours)^2 −0.06, 95% CI −0.11 to −0.01] and life

satisfaction [linear relationship; time outdoors (hours) 0.06, 95% CI 0.04 to 0.07]. There was no evidence that time outdoors mediated the relationship between EVI and WEMWBS/life satisfaction. EVI was not associated with CMD (EVI 0.4 to 0.8 vs. < 0.2, AOR 1.06, 95% CI 0.80 to 1.42) but increased time outdoors was significantly associated with reduced odds of seeking help for CMD (AOR 0.96, 95% CI 0.93 to 0.99). Proximity of nearest GBS was not related to WEMWBS/CMD/life satisfaction.

Research question 5: Are associations between residential GBS exposure and access to GBS and use, and subjective well-being mediated by walking-specific physical activity and modified by socioeconomic disadvantage?

Prerequisite conditions were met for potential mediation, but the time outdoors-WEMWBS coefficient was not substantially attenuated when adding walking physical activity (standardised time outdoors beta reduction: 0.21 to 0.19, standardised time outdoors^2 beta reduction 0). However, the time outdoors-life satisfaction coefficient was partly attenuated when adding walking physical activity (beta reduction 0.07 to 0.05). The 95% CI before and after addition of walking physical activity overlapped, evidencing that walking physical activity partially mediated the relationship between time outdoors and life satisfaction.

We found no evidence that material deprivation modifies the association between EVI/GBS proximity and WEMWBS/life satisfaction (WEMWBS, EVI × in material deprivation vs. EVI × not in material deprivation: 3.24, 95% CI −16.90 to 23.38; EVI^2 × in material deprivation vs. EVI^2 × not in material deprivation: −5.31, 95% CI −33.48 to 22.85). For those in material deprivation, the association between leisure time outdoors and both WEMWBS and life satisfaction was stronger compared with those not in material deprivation up to ~4 hours leisure time outdoors/week (7 + hours for life satisfaction) (WEMWBS: time outdoors × in material deprivation: 1.41, 95% CI 0.39 to 2.43; time outdoors^2 × in material deprivation −0.18, 95% CI −0.33 to −0.04; life satisfaction: time outdoors × in material deprivation: 0.08, 95% CI 0.03 to 0.13).

Increasing time outdoors had a greater benefit (higher WEMWBS/better life satisfaction) although well-being remained generally lower.

Public involvement

We co-created our study analysis plans and interpreted our findings with stakeholders. We successfully engaged policy makers but were less successful with third-sector mental health stakeholder engagement on our steering group. This was due to staff changes restricting opportunities for building relationships and increasing pressures on those working in mental health due to the COVID-19 pandemic.

Conclusions

Despite some limitations due to a lack of EVI variation in Wales, the findings from our study considerably extend the evidence on the association between mental health/well-being and residential GBS/visits. Particularly important are our findings of a disproportionate well-being benefit of time outdoors for individuals living in material deprivation.

We saw an increase in mental health help-seeking behaviour with greater EVI for people in the most deprived rural areas, potentially highlighting isolation as an issue. Conversely, for those with an historical CMD, greater EVI may reduce mental health inequalities.

Neighbourhood greenness and visiting GBS were both related to better well-being, but the lack of mediation suggests different mechanisms.

There is a need to understand better where to prioritise investment for equitable availability of GBS or the targeted delivery of interventions to reduce health inequalities.

Research recommendations

More sophisticated methods are needed to spatially impute changes in GBS recorded in planning data; current methods are restricted by data capture timelines.

1. The individual-level linked survey and routine health data can provide baseline data for future prospective longitudinal analysis, and more sophisticated studies of mediation and moderation in environment-health research.
2. Consent for individual level data linkage should be considered for all surveys; the depth of behavioural insight nested within routinely collected data for a population provides invaluable insights into explanatory mechanisms.
3. Further analyses could investigate different 'settling in' periods following moving home, capturing outcomes at the time of the move and subsequently.
4. The dataset could be extended to investigate maternal mental health outcomes and general outcomes for children.

Funding

Chapter 1 Green and blue space and mental health

Background and aims

There is growing evidence that green and blue outdoor spaces positively impact mental health and well-being. Impact occurs through mechanisms such as increased social contact, cognitive improvement, reduced stress and psychological recovery.[1–5] Green and blue space (GBS) may also mitigate environmental stressors such as urban noise and poor air quality.[6,7] Currently, there is limited evidence for a causal relationship between the 'greenness' and 'blueness' of environments and adult mental health;[5,8] a key limitation being the lack of longitudinal studies.[9] We present the largest, most comprehensive, longitudinal evaluation of the impact of changes in exposure to GBS on mental health and well-being among adults in Wales. We conducted this study in Wales due to the mechanism allowing dynamic linkage of environmental metrics to individual's health data for the population of Wales, within the Secure Anonymised Information Linkage (SAIL) databank.[10]

Mental health and well-being

Mental health is defined by the World Health Organization (WHO) as 'a state of well-being in which every individual realizes his or her own potential, can cope with the normal stresses of life, can work productively and fruitfully, and is able to make a contribution to her or his community'.[11] Wellbeing is a complex multidimensional concept typically considered to evaluate whether or not life is worthwhile and has purpose.[12] Subjective well-being is an important marker of quality of life, and is related to mental and physical health outcomes, including survival.[13–15]

Mental ill-health is one of the main contributors to the non-fatal global disease burden. It accounts for between 21.2% and 32.4% of years lived with disability,[16,17] and impacts all age groups.[18] In the UK, the Adult Psychiatric Morbidity Survey found that more than one-fifth of adults (aged 16 years or over) live with a common mental health problem.[19] In addition, almost half (43.4%) of UK adults report a diagnosable mental health condition at some point in their life, with rates higher in women than men (51.2% and 35.2%, respectively).[19] The 2015 Welsh Health Survey found that 13% of Welsh adults (aged 16 years or over) were currently receiving treatment for a mental health disorder.[20]

Mental ill-health poses a significant burden on health and social care systems. It costs the UK economy an estimated £70–100 billion per year, around 4.5% of the gross domestic product.[21] In Wales, over 2007–08, the costs of mental ill-health (£7.2 billion) outweighed the health and social care costs of all other forms of illness (£6.1 billion).[21] In this study we refer to common mental health disorders (CMDs) that comprise anxiety and depressive conditions that are generally treated by a general practitioner (GP).

Green and blue spaces

Blue spaces are dominated by a watery element, such as a lakeside, river or coast. Green spaces may include a watery element but are characterised by predominantly 'green' elements such as grass or trees. Green spaces, and the networks of those spaces, are defined by Natural Resources Wales (NRW)[22] as a:

> shorthand for both vegetated land-parks, community gardens, trees, woodlands and hedges, informal spaces, allotments and food growing sites, areas of water, such as rivers, canals, lakes and ponds and the sea shore. Green infrastructure describes the network created by all these spaces.

Green or blue spaces maybe be highly managed, such as urban parks, or more 'natural' spaces, such as native woodland. There is not necessarily an assumption of ownership and access rights; typically, this will be defined if relevant (e.g. 'publicly accessible green spaces' could refer to an urban park). The language used to describe environmental specifics varies but without a discernible pattern according to discipline or context.[23] Terminology also varies according to the intended 'use' of a space or resource. We use the terms 'green space' and 'blue space', defined as spaces dominated by natural elements, in either rural or urban contexts, with green and blue space hereafter abbreviated 'GBS'. *Chapter 3* presents the comprehensive, spatially nested typology of GBS developed for this study.

Green and blue space and mental health and well-being

Green space

Reviews have shown that exposure to green space can promote life satisfaction,[24] aid psychological restoration[25] and mitigate the impact of stress.[26] Views of green space from the home have been linked to measures of subjective well-being, and there are relationships between positive mental health in adulthood and overall greenness of the living environment,[5] residential proximity to GBS[27,28] and frequency of green space visits.[29,30] However, these relationships are not necessarily straightforward. The complexity of characterising exposure to green spaces and the pathways through which they affect mental health means that causal relationships are not always clear.[8,31] Most caution that the evidence is currently limited and often inconsistent.[8]

Factors such as age, sex, physical activity, ethnicity and socioeconomic status appear to moderate relationships between exposure to green space and mental health outcomes in adulthood.[27,32,33] The associations between greener living environments and mental health appear to vary over the life course[34] and there is little consistency according to gender.[35] Green space may be of greater benefit to people from disadvantaged groups and living in more socioeconomically deprived areas compared with those in less deprived areas.[8,27] Again, this is not a straightforward relationship and is not consistently statistically significant in studies investigating the effects of socioeconomic status.[31] Less is known about interactions between ethnicity and the salutogenic effects of green space exposure, but a UK birth cohort study showed that ethnicity moderated relationships between residential green space and mental well-being.[33]

Blue space

The hazardous qualities of blue spaces are well known, but their salutogenic public health implications have, until recently, received less scientific attention (and less than those of green space).[36,37] There is a growing body of evidence that exposure and proximity to blue space is linked to positive mental health and well-being.[38] Living near the coast is shown to have general mental health benefits.[4,39] There is also evidence that people living near (or with views of) the coast experience lower psychological distress.[40,41] Epidemiological studies do not typically distinguish between coastal and freshwaters,[41] but there is some evidence that proximity to inland water also yields positive mental health effects.[42,43]

Mediators and moderators of relationships between blue space and mental health are less well understood than those for green space, but it is likely that similar factors apply. For example, a 2019 study in England showed that the mental health benefits of residential coastal proximity appear to moderate the effects of low socioeconomic status.[44] Physical activity also appears to be a mediator between positive mental health and coastal proximity. This effect is limited to terrestrial physical activity near the coast rather than 'watersports'.[43]

Quantity, proximity, quality and type of green and blue space and mental health and well-being

The amount of green space in an area has been linked to good population mental health[4,45] and reduced inequalities in health in the UK.[44] However, there is insufficient evidence to dissect how

type, proximity, quantity and 'quality' (either ecological or in terms of maintenance) or 'qualities' (the characteristics) of green or blue space each influence population-level mental health and well-being.[46] Most available evidence linking GBS to health outcomes considers relationships at a local area level, typically the total amount of green space around residences, with no assumption about type, quality or usage.[5]

The type, quality, and qualities of GBS may be significant in the context of mental health benefits. Individual studies suggest that environment (or landcover) type is an important consideration for general and mental health.[3,39,47,48] In urban settings, gardens and allotments are notably beneficial for mental health.[49,50] The impact differs according to the type of garden, with private spaces appearing to be most strongly associated with better outcomes.[51] In other studies, proximity to specific types of environments, such as coasts, have been found to be linked to better mental health.[44]

The impacts of GBS 'quality' on mental health and well-being are more difficult to assess, given the breadth of meaning that can be attached to this term. We consider 'quality' as the state, whether ecological or physical, of the green or blue space. 'Qualities' refers to the characteristics, or composition of the space. GBS 'quality' and 'qualities' are multidimensional concepts. They may be defined and assessed subjectively (e.g. perceived ecological quality, perceived safety) or objectively (e.g. public facilities, air pollution levels) in a number of ways. Evidence of links between indicators of GBS quality and mental health is limited.[27,52] However, there are some studies supporting 'quality' and 'qualities' of GBS as important components for mental health and well-being impacts. General satisfaction with local green space is important,[33,53] and perceptions of safety and maintenance determine people's use of green space.[54,55] Public accessibility and related urban planning influence the likelihood of the use and resulting experience of spaces, impacting potential health gain.[53,56] The relative biodiversity of environments may also play a role, particularly in terms of the types of nature people experience,[57,58] but evidence is limited and inconsistent.[59]

Overall, the type, proximity, quality, and qualities of GBS are likely to be important in the context of mental health benefits. However, current evidence is inconsistent for relationships between the: quantity,[5] type[60] and quality and qualities of GBS,[27] and mental health and well-being.

Use of green and blue spaces

The studies mentioned previously often make no assumption about usage of the GBS. Usage of a space is theorised to be a factor in potential benefit;[61] however, actual visits to, and time spent in, GBS are less well studied.[5] Studies have, however, shown that recreational use of GBS is associated with multiple mental health outcomes.[62,63] Further, there are indications that minimum weekly baselines of time spent in green spaces are linked to well-being and good mental health.[63] Whether or not physical activity mediates benefit has been investigated in a number of studies, with mixed results.[64]

System-wide impacts of green and blue spaces

Additional pathways may link the presence of GBS to mental health. These include the mitigation or avoidance of the effects of urban heat islands, air and noise pollution, and flooding.[46] Although evidence is limited in terms of demonstrable associations of these pathways with better mental health,[65] some linkages have been demonstrated; for example, green spaces were found to be associated with reduced psychological distress from local noise pollution.[66] Thus, there are likely to be multiple benefits of GBS, impacting our physical and mental health through indirect mechanisms within wider ecological systems.

Inequalities

There are inequalities in the distribution and accessibility of good quality GBS in the UK. People in deprived areas, minority ethnic communities, elderly people and those with long term poor health and disabilities often have less (physical) access to good quality green spaces. They tend to use green spaces less and are more likely to have negative perceptions regarding their usage of such spaces.[67–69]

These groups also tend to have poorer health outcomes.[70,71] However, there is cross-sectional evidence suggesting that health inequalities are lessened with greater proximity to green or blue spaces.[44,72,73]

Policy context: green and blue space and mental health and well-being

Addressing the causes of poor mental health, promoting solutions and reducing burden to individuals, communities and society is a key concern for the UK and Welsh governments. Where we live, learn, work and play are determinants of mental health and influence key risk factors.[74] As such, the relevance of place to mental health is recognised across contemporary policy and service delivery strategy in Wales, as it is across the UK, and is embedded within decision-making frameworks such as the Well-being of Future Generations (Wales) Act.[75]

The Healthier Wales strategy notes the importance of living and working conditions on health outcomes.[76] The strategy details a 'whole systems' approach to health promotion and protection in Wales, recognising that health services are only one element of what determines health. This systems approach is reflected in the Well-being of Future Generations (Wales) Act.[75] The Act requires public bodies, including the health and environmental services, to work together to address health, inequalities and living conditions. It also identifies a set of 'well-being' goals that all public bodies must work to achieve and requires these bodies to consider the long-term impacts of their decisions, including how these relate to the well-being of people in Wales. The Act and associated guidance highlight the need to ensure equitable provision of good quality living environments, including green and biodiverse spaces, for health. The decision-making frameworks associated with the delivery of the Act, especially the public services boards, are intended to reflect these interdependencies, ensuring that health impacts of decisions relating to the environment, and vice versa, are fully considered.

The importance of the natural environment, and green or blue spaces, in meeting these well-being goals is reflected in national policy. For example, the Welsh Together for Mental Health strategy states that natural environment can have significant positive impact on physical and mental well-being, noting that access is positively related to a number of mental health, developmental and behavioural outcomes and risk factors.[75] *Creating an Active Wales*[77] highlights the role of green space provision in supporting population physical activity.

The importance of place, and particularly green or blue spaces, to health is similarly reflected in the policies and strategies of NRW. NRW's *Outdoor Recreation and Access Enabling Plan 2015–2020* details the health benefits of outdoor recreation.[22] The plan aims to increase appreciation and use of the natural environment to improve people's health and well-being. They aim to deliver impact by ensuring equitable access to spaces, promotional and engagement activities, and through partnership working such as with the planning services. NRW's *State of Natural Resources Report* 2016 suggests increasing green infrastructure (trees and green space) and increasing woodland cover as key for sustainable management of natural resources and the well-being of people in Wales.[78] NRW's *Managing Today's Natural Resources for Tomorrow's Generations*[79] includes the aim 'to make better use of local green space in both urban and rural areas the norm, together with a recognition that physical activity in the outdoors contributes to the prevention of many physical and mental illnesses'. Proximity to accessible green space and physical activity in the outdoors and its link to physical and mental health are identified as key indicators. NRW's *Improving Access for All* policy acknowledges and pledges to tackle potential inequalities in access to GBS.[80]

In Welsh planning policy *Facilitating Accessible and Healthy Environments* is one of the five key national sustainable placemaking outcomes and has the specific ambition to provide 'accessible and high-quality green space'.[76] The strategy emphasises the role of GBS within wider networks of green infrastructure. It highlights the potential to enhance health and well-being through providing a context for physical activity, higher-quality living environments and quiet tranquil areas.

4

Evidence gaps and requirements

Decision makers at all scales, and from multiple sectors (public and third especially), are seeking effective evidence-informed ways to provide, manage and promote GBS for health. England's Department for Environment, Food and Rural Affairs, for example, recently committed to investing in research to inform their service delivery, and in the provision and use of natural environments for health outcomes.[81] NRW's first *State of Natural Resources Report 2016* explicitly linked Wales's natural resources to the well-being of people in Wales.[78,79]

Improved evidence on the relevance of green or blue spaces to mental health is required to inform planning and area regeneration in the UK, which shape the environments people are exposed to. While there is a substantial body of evidence indicating the benefits of GBS for mental health and well-being, there are limitations to the existing evidence-base. There are high levels of heterogeneity between studies, in measures of GBS exposure and mental health and well-being outcomes, limiting the potential for robust meta-analyses.[82] There is also a lack of evidence relating to certain population subgroups and potential pathways. There is insufficient evidence for determining the most appropriate size, location, configuration, connectivity, composition and characteristics of GBS for mental health outcomes.[82]

There are few longitudinal studies on either change in GBS exposure though time or on the impact of visiting GBS over the life course. In cross-sectional studies, which currently provide much of the evidence on the relationships between GBS and mental health and well-being, the direction of relationships is not always clear. For example, cross-sectional studies may be unable to determine whether green space exposure leads to good mental health, or if people with good mental health visit green space more often. Longitudinal studies enable a sequence of events to be constructed (e.g. mental health measured prior to and following a change in exposure), allowing for more confident inferences regarding causality.[83] The small number of studies using longitudinal data are beginning to help unpick the causal direction of impact.[45,84]

There is evidence that public health interventions are cost effective in reducing health service expenditure,[85] but evidence relating to nature-based interventions is limited,[86,87] particularly in the context of mental health. Systematic synthesis of nature-based interventions is difficult due to plurality in valuation methodologies, but emerging evidence suggests significant value to health and social systems.[86] Urban green spaces in the United States have been estimated to be worth between US$2.7 and $6.8 billion annually (2012).[88] Physical activity taking place in natural environments in England has been estimated to be worth £2.18 billion in welfare gains,[89] with a substantial proportion (£176 million) potentially derived from activity in marine environments.[90] Natural areas must be accessible for such benefits to be gained. Pedestrian infrastructure, such as the Wales Coast Path, encourages physical activity leading to health benefits with high economic value, estimated at £18.3 million annually.[91]

Aim and objectives of the project

We aim to create novel linkages between environment and health data to enable quantification of the impact of exposure to GBS through time for a national population.

This project will address evidence gaps by:

- providing longitudinal measures that factor in loss and enhancement of exposure to GBS
- explicitly considering population subgroups and health inequalities
- creating exposure measures that include a variety of GBS characteristics, allowing planners to consider different configurations of spaces and their potential benefits
- modelling the impact of GBS exposure on general practice events (GPE) and associated NHS costs.

Our project will also contribute to addressing the evidence needed by our project partner, NRW, for 'place-based assessment of potential changes, both positive and negative, to the "total health impact" (mental, physical and well-being) arising from future land management options'.[92] We will also produce findings to underpin future interventions, for example estimating effect sizes for interventions aiming to increase frequency of visits to GBS.

We deliver this project through a series of objectives, using data from a national population, and a nested survey sample.

National population dataset objectives

- Create a longitudinal dataset of residential GBS exposure data for all homes in Wales using UK Ordnance Survey (OS), local authority and remotely sensed satellite data.
- Create an 11-year dynamic cohort of individual-level longitudinal residential GBS exposure to answer the questions: 'Is a greater residential exposure and access to GBS associated with the likelihood of seeking help for a CMD in general practice?' and 'Is the association between exposure, or access to, GBS and seeking help for a CMD modified by socioeconomic disadvantage and moving home?'

Survey dataset objectives

- Create longitudinal, environment, health and demographic individual-level data linkages between survey and routine data within the SAIL databank.
- Utilise the data-linked National Survey for Wales (NSW) to answer the question: 'What is the impact of GBS residential exposure, modified by GBS use and multiple socio-physical modifiers, on well-being and CMD?'

The national population dataset uses routinely collected health and demographic data (data collected for purposes other than research). The survey sample dataset links routine data in the SAIL databank with in-depth survey responses from the NSW. Both consider socioeconomic health and well-being inequalities.

Synthesising objective
To translate results into policy recommendations for government and disseminate results from health outcome and economic impact assessment to academics, stakeholders, and policy makers.

Chapter 2 Research design and cohort description

Introduction

In this chapter we summarise the research design, the methods used to create our cohort, and the socio-demographics of the cohort. We outline key exposure and outcome variables together with the quantitative data analysis methodology. Detailed descriptions of the variables and analytic approaches used for different aspects of the study are described in subsequent chapters.

Research design

We performed a retrospective, population-wide study. We constructed the cohort using data from the Welsh Demographic Service Dataset (WDSD) and linked with residence-level environmental metrics (derived from satellite imagery and planning data) and longitudinal, routinely collected, population-scale, individual-level, anonymised electronic health record (EHR) data from the NHS. The cohort had individual-level linkage for a subgroup who were surveyed (cross-sectionally) by the NSW from Welsh Government. The survey included questions on well-being and visits to outdoor spaces.

The study aimed to quantify the impact of access to GBS and ambient greenness around the home on CMD and SWB for the national population of Wales. This project examined the following research questions:

1. Is a greater residential exposure and access to GBS associated with the likelihood of seeking help for a CMD in general practice?
2. Is the association between exposure, or access to, GBS and seeking help for a CMD modified by socioeconomic disadvantage and moving home?
3. Is a greater residential exposure associated with the likelihood of using general practice services?
4. Is greater residential GBS exposure and access to GBS associated with higher SWB, or a lower likelihood of seeking help for CMD, and are these associations mediated by intentional use of GBS for leisure?
5. Are associations between residential GBS exposure and access to GBS and use, and SWB mediated by walking-specific PA, and modified by socioeconomic disadvantage?

We used the total adult population cohort to address questions 1–3 (see *Chapter 5* and *Chapter 6*) and a representative sample of the adult population in Wales based on the NSW for two years (2016–17 and 2018–19) to answer questions 4 and 5 (see *Chapter 7*).

Ethics considerations

We used data available in the SAIL databank at Swansea University, Swansea, UK.[93,94] All proposals to use SAIL data are reviewed by an independent information governance review panel. The panel considers each project to ensure proper and appropriate use of SAIL data. If access is approved, it is gained through the SAIL gateway, a privacy-protecting safe haven and remote access system. This study was approved by the panel as project 0562.

Cohort creation

The cohort was created using data from the WDSD linked to:

- Annual District Death Extract: mortality information including the date and cause(s) of death from the Office for National Statistics (ONS) mortality register.
- Patient Episode Database for Wales: demographic and clinical data on all inpatient and day case admissions to NHS Wales hospitals and Welsh residents treated in other UK nations.
- Welsh Longitudinal General Practice (WLGP): information on symptoms, diagnoses, prescriptions and referrals collated from clinical information systems in use at each general practice in Wales.
- Welsh Index of Multiple Deprivation (WIMD, 2011 and 2014), the Welsh Government's official measure of relative deprivation for small areas in Wales.[95]
- Rural–urban ONS classifications (2011) at lower-layer super output area (LSOA).[96]
- NSW, an annual, repeat cross-sectional survey of about 12,000 adults in Wales.

The cohort, designated the GBS e-cohort, is held in the SAIL databank, a privacy-protecting trusted research environment providing secure storage of anonymised, person-based demographic, health, social and education data for the population of Wales.[93,94]

Environmental exposures

Chapter 4 describes environmental exposures. In brief, we derived measures of ambient residential greenness exposure and potential access to GBS for all residences in Wales using:

1. Satellite data (Landsat™ 2008–2019)[97] to create annual greenness densities of mean enhanced vegetation index (EVI) and normalised difference vegetation index (NDVI) within 300 m of each residence.
2. OS MasterMap (OSMM) Topography Layer[98] (2018) to capture natural and man-made features, including the outline of homes and parks.
3. OSMM-derived Greenspace dataset (2018).[99]
4. Local authority technical advice notes (TAN 16), legally required records of data on sport, recreation and open spaces managed by local authorities.[100]
5. Open-source portal data from the Lle Geo-Portal (forestry, urban tree cover).[21]
6. OpenStreetMap road/footpath data.[101]

Ambient greenness within 300 m of each residence was measured yearly and linked to the cohort in three-component variables that represent ambient residential green/blueness:

- Mean EVI (minimum, mean, median, max)
- Mean NDVI (minimum, mean, median, max)
- Coastal and/or inland water (yes/no)

The potential for an individual to access a range of types of GBS along a network of paths and roads within 1600 m of each home was modelled for 2018. Access measures have multiple variables available at three (nested) levels of detail, following consultation with stakeholders (see Chapter 3):

- Level 1 divides green spaces into amenities, functional spaces, seminatural habitats, enclosed and linear spaces.
- Level 2 provides more detail on the type of space; for example, identifying a recreation space.
- Level 3 contains greater detail; for example, identifying play areas within a recreational space.

Ambient residential greenness, and potential to access GBS, were augmented by survey responses about leisure time visits to outdoor spaces in Wales for the NSW subgroup.

Changes in access and exposure to GBS could be either at the environment level (e.g. a change in the GBS itself), or at the individual level, such as through moving home.[83]

Survey data

A cohort subgroup responded to NRW questions in the 2016–17 and 2018–19 NSW.[102] The NSW is an annual repeat cross-sectional, government-sponsored omnibus survey of a representative sample of the population of Wales (annual *n* ~12,000). Topics include education, culture, health and well-being, and more detailed information on socioeconomic circumstances than found in administrative data. The NRW questions[102] ask whether respondents visited outdoor spaces in Wales, including time spent outdoors on leisure activities, and types of activities undertaken. The NSW respondents aged 16 years and over who consented to NSW-administrative data linkage (> 90%) were linked to the cohort.

Mental health outcomes

Common mental health disorders

A count of specific GPE relating to CMD (anxiety and depression, see *Report Supplementary Material 1* for read codes), were extracted from WLGP and aggregated quarterly. The WLGP is collated from clinical information systems at each general practice around Wales and uses read codes recorded during a GP consultation. Read codes are a hierarchical nomenclature used to record clinical, referral and administrative information. Test results are electronically transferred into the WLGP from secondary care systems.

To identify people seeking care in general practice for CMD, we applied an existing algorithm designed to estimate CMD prevalence (anxiety and depression) with high specificity. The algorithm is based on diagnoses, symptoms and treatments (see *Report Supplementary Material 1*).[33] In validating the algorithm, the authors compared routine data results with those from a Mental Health Inventory-5 (MHI-5) survey. The authors suggested several different algorithms providing different levels of sensitivity and specificity. We chose the algorithm with the largest sensitivity without compromising on specificity. The authors concluded that their algorithm may be used to create outcome measures for trials and cohort studies.[103]

Using the algorithm, we identified people with CMD each quarter based on treatment of historical diagnosis(es), and/or current diagnoses or symptoms (treated or untreated) recorded in read codes (see *Report Supplementary Material 1*) in their WLGP data (algorithm 10).[33] The algorithm identifies 'current' diagnoses/symptoms as relevant read codes in the preceding one-year period. It identifies 'historical' diagnoses through a search for relevant read codes outside the 'current' period. The length of retrospective data for each individual depends on the length of their registration with a general practice supplying data to SAIL.

CMD treatment was defined as at least one prescription for an antidepressant, anxiolytic or hypnotic in the 1-year current period.[1] We did not include non-drug treatments in our CMD case definition as this information is not available in the WLGP dataset. The algorithm required a diagnosis or symptom of a CMD prior to counting treatments. This is because several treatments have multiple purposes outside mental health disorders (e.g. antidepressants for pain treatment). In common with previous studies, we excluded the more severe condition of psychosis from these 'common' disorders to maximise the potential to detect change due to GBS exposure.[104]

We identified adults (16 years and over) seeking GP care for a CMD in the WLGP data. We acknowledge that community prevalence will be significantly greater than seen in primary care data, because only about one-third of people affected by CMD seek help in primary care.[105]

Subjective well-being

The NSW uses the Warwick–Edinburgh Mental Well-being Scale (WEMWBS) to capture affective and evaluative well-being.[106] Participants are presented with a series of 14 positively worded statements regarding their thoughts or feelings. WEMWBS scores are designed to assess population level mental well-being.[106]

Life satisfaction

In the NSW, life satisfaction is a measure of evaluative well-being based on responses to the question 'Overall, how satisfied are you with your life nowadays?'.[107] This gives a longer-term view of an individual's mental well-being than WEMWBS.

General practice events

A daily count of all GPE per person (e.g. appointments, tests), not limited to those relating to CMD, were extracted from WLGP and aggregated quarterly. This eliminated counting multiple test results. These events are not only mental health-related, so reflect a general need for primary health-care services.

Data linkage

A strength of the SAIL trusted research environment is the split-file process for anonymising all individuals and households in Wales. This overcomes confidentiality and disclosure issues arising in health-related data warehousing.[93] The data sources described above were split at the source organisation into demographic (identifiable) data and clinical (non-identifiable) data. A temporary system linking field ensures the data can be rejoined later in SAIL.[93] The identifiable component was sent to our trusted third party, Digital Health and Care Wales, who anonymised and encrypted the identifiable data, assigning a unique anonymised linking field (ALF), the primary key variable for record linkage, and a residential ALF (RALF) based on the unique property reference number of their place of residence. Data were rejoined in SAIL using the temporary system-linking field, and the ALF was further encrypted before being made available to project statisticians. Detailed information on the matching technique used to assign an ALF to each individual across diverse administrative datasets has been published elsewhere.[93] Briefly, a matching algorithm is applied in sequence. Records with valid NHS numbers are accepted, deterministic record linkage is then carried out on a set of matching variables and, finally, the probabilistic record linkage is carried out on the remaining unmatched records.[93]

The environmental exposure indices created within a secure geographic information system (GIS) platform (the Secure eResearch Platform) were linked at household level with demographics from WDSD and NSW using the RALF for record linkage. This approach has been successful in a range of public health studies.[10,108–110] The dates of each change of address are recorded in the WDSD, providing duration of residency and the ability to link our time-varying environmental exposures to each residence and LSOA as individuals move home.

WDSD contains dates of change of address provided by patients to their GP. We aggregated these data quarterly, so individuals can have multiple RALFs within a three-month period. We selected one RALF per quarter using detailed rules to identify and rank RALFs (*Appendix 1*, including *Table 22*) to assign one set of environmental metrics to an individual each quarter. Briefly, a RALF confidence score (low/intermediate/high confidence in address match) is provided within the WDSD. To determine the most highly ranked RALF within the quarter, this measure of linkage confidence is combined with the total duration (start date to end date) within that RALF. RALFs with higher linkage confidence and which are assigned to the individual for longer are given priority over RALFs with lower confidence and shorter durations. When more than six RALFs are linked to an individual within a three-month period, the quality of data is considered unreliable, and no RALF is assigned during that period.

Cohort variables

The variables available for the main cohort include: (1) sociodemographic and economic characteristics; (2) CMDs/well-being; (3) comorbidity index; (4) social environment and life events (births/deaths in the household); (5) environmental metrics; and (6) other administrative cohort information (*Table 1*). NSW variables are described in *Chapter 7*.

TABLE 1 Cohort variables

Domain	Subdomain	Individual (I)/ residence (R) level
i. Sociodemographic and economic characteristics	Age	I
	Sex	I
	Deprivation*	R
	Urban/rural	R
ii. CMDs/well-being	Depression	I
	Anxiety	I
	CMD	I
iii. Comorbidity index/hospital episode count	Modified Charlson Comorbidity Index+	I
	Inpatient hospital episode i	I
iv. Social environment and life events	Birth in household	R
	Death in household	R
	Household composition (count of children < 16 years in household)	R
	Time since last residential move	I
v. Environmental metrics	EVI	R
	NDVI	R
	Access to GBS (distance/size/ type)	R
vi. Other administrative cohort information	Cohort entry/exit reason (death/ migration)/date	I
	ALF ii	I
	RALF with from/to dates ii	R
	LSOA	R

* 2011 and 2014 WIMD as defined by the Welsh Index of Multiple Deprivation (IMD) quintiles 2011 and 2014;[29] plus Charlson Comorbidity Index as defined by the NHS Information Centre. (i) inpatient hospital episode as identified in the Patient Episode Database for Wales; (ii) ALF and RALF are individual and household ALFs, respectively within the SAIL Databank.

Analysis platform

The SAIL databank operates on the UK Secure Research Platform.[110] All analyses were conducted in RStudio (version 1.4.1103) with R (version 4.1) and MLwiN (version 3.05) through a remote desktop protocol within the SAIL gateway.[111]

Cohort description

The cohort inclusion criteria were all persons aged 16 years and over registered with a GP providing patient records to the SAIL databank between January 2008 and October 2019, giving a cohort of 2,800,483 individuals. Exclusion criteria were not being registered with a GP providing data to the SAIL Databank, not having a Welsh residential address between January 2008 and October 2019, and not having sex or week of birth recorded in WDSD. The WDSD contains demographic characteristics of people registered with a free-at-the-point-of-care GP in Wales providing data to the SAIL databank (80% population coverage).[112]

We intentionally removed people who did not fit with study criteria from the total people recorded in WDSD (*Figure 1*). This comprised those who did not reach the age of 16 years during the study period, who died or moved to a home outside of Wales before January 2008 or moved out of Wales before turning 16 during the study period. Additionally, we excluded those missing key demographic data (sex, week of birth; *n* = 12,210). We further excluded 839,063 individuals who had missing data; for example, they were not registered with a GP providing data to the SAIL databank, did not have a Welsh residential address between January 2008 and October 2019 or did not have sex or week of birth recorded in WDSD. Those who meet the criteria after the study start are censored until their qualifying entry date and those who exit before the study finishes (through death or migration) are right censored.

Responses to NRW questions in the 2016–17 and 2018–19 NSW were available for a cohort subgroup of 5312 individuals.

The cohort has 24.9 million person years of follow-up. An additional average of 30,238 people joined the cohort annually through migration into Wales, and an average of 34,709 people by reaching age 16 years, totalling 710,570 (25%). Annually, an average of 22,987 people died and 1603 moved out of Wales permanently, totalling 294,437 (10.5%). The cohort sample was broadly representative of the Welsh population compared with 2011 Census figures (*Table 2*).[113] More than one-fifth (22.6%) of the
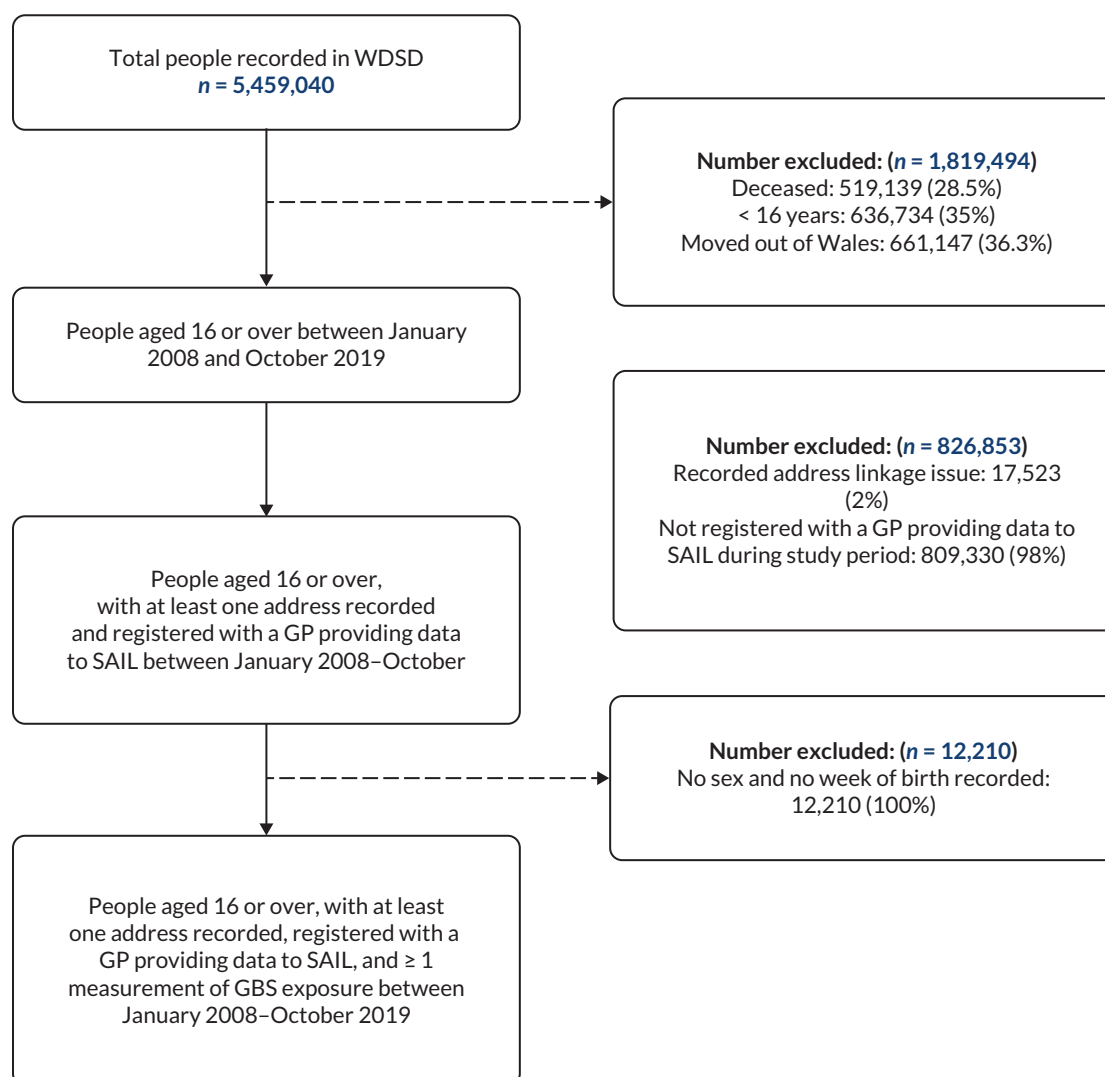
FIGURE 1 Cohort construction of the 2,800,483 individuals using the demographic dataset (WDSD) and linkage to the WLGP dataset.

TABLE 2 Sociodemographic characteristics of the cohort at baseline

Sociodemographic characteristic	Whole cohort			Moved home at least once		
	(n)	(%)	Difference relative to reference group	(n)	(%)	Difference relative to reference group
Sex						
Male	1,381,576	49.1	−2.8	561,868	47.2	−10.5
Female (ref)	1,419,907	50.5	–	628,034	52.8	–
Age group						
16–21	614,265	21.8	473.7	316,803	26.6	696.1
22–30	418,046	14.9	292.1	264,988	22.3	565.9
31–40	405,533	14.1	271.1	201,099	16.9	405.3
41–50	409,772	14.6	284.2	149,919	12.6	276.7
51–60	353,182	12.6	231.6	101,296	8.5	154.5
61–70	303,247	10.8	184.2	68,420	5.8	71.9
71–80	190,964	6.8	78.9	47,581	4.0	19.6
80+ (ref)	106,482	3.8	–	39,796	3.3	–
Rurality						
Town and fringe	452,951	16.7	351.7	778,507	69.9	405.1
Urban > 10K	1,847,233	68.2	10.6	181,507	16.3	17.8
Village, hamlet & isolated dwellings (ref)	408,559	15.1	–	154,125	13.8	–
WIMD quintiles						
1 – Most deprived	568,394	20.8	3.0	254,944	22.6	21.2
2	544,315	19.9	−1.5	229,384	20.4	9.1
3	559,434	20.5	1.5	226,951	20.1	7.9
4	508,838	18.6	−7.9	205,130	18.2	−2.5
5 – Least deprived (ref)	552,939	20.2	–	210,323	18.7	–
Home moves 2008–19						
Does not move	1,611,581	57.5	183.3	–	–	–
Moves once	622,025	22.2	9.4	–	–	–
Moves more than once	567,877	20.3	–	–	–	–

Baseline is defined as the first period an individual enters the cohort. Relative difference is calculated as the percentage difference to the reference group. Rurality defined using ONS settlement type categories.[41]

adult population in the most deprived quintile moved home at least once during the cohort period 2008–19, with fewer moving in the least deprived (18.7%) and next least deprived (18.2%) quintiles (*Table 2*). Younger people (< 30 years) and those living in the most deprived areas had the highest prevalence of moving at least once during their time in the cohort (44.9% and 22.6%, respectively; *Table 2*). To check for potential bias, area-level characteristics of residences linked and unlinked to the cohort were compared. The percentage of unlinked homes did not increase with deprivation. However, a higher proportion of unlinked homes were in rural areas (*Table 3*).

TABLE 3 Comparison of linked and unlinked homes

	All		Linked to cohort		Not linked	
	(n)	(%)	(n)	(%)	(n)	(%)
WIMD quintiles						
Most deprived	292,733	19.5	243,928	20.7	48,805	15.3
Next most deprived	302,100	20.2	248,265	21.0	53,835	16.9
Mid deprived	315,169	21.0	241,919	20.5	73,250	23.0
Next least deprived	309,795	20.7	219,215	18.6	90,580	28.5
Least deprived	278,323	18.6	226,490	19.2	51,833	16.3
Rurality						
Rural town and fringe	197,499	13.2	161,417	13.7	36,082	11.3
Rural town and fringe in a sparse setting	69,875	4.7	42,346	3.6	27,529	8.6
Rural village and dispersed	101,978	6.8	70,118	5.9	31,860	10.0
Rural village and dispersed in a sparse setting	127,178	8.5	80,361	6.8	46,817	14.7
Urban city and town	973,872	65.0	802,972	68.1	170,900	53.7
Urban city and town in a sparse setting	27,718	1.9	22,603	1.9	5,115	1.6

Rurality defined using ONS settlement type categories.[41] WIMD quintiles.[95]

Conclusion

We constructed a cohort of the population of Wales between January 2008 and October 2019 using data from the WDSD, household-level ambient greenness from satellites, access to GBS from planning records, WLGP data on CMD and service use, and for a subgroup, the NSW (2016/17 and 2018/19) in the SAIL databank. The cohort included 2,801,483 individuals (99,682,902 observations). In the following chapters we describe how we used this cohort to quantify the impact of ambient greenness and access to GBS on CMD and subjective well-being for the national population of Wales.

Chapter 3 Developing a typology of green and blue spaces

Introduction

There is no universal definition of GBS;[114] definitions are often nuanced, context-specific, and application dependent.[114] Heterogeneity in GBS definitions and how they are operationalised can present challenges to understanding the relevance of the resource for health and well-being outcomes, and for translating evidence into policy and action. Epidemiology and public health research often assess GBS using vegetation indices based on data from satellite imagery, quantity (area) of GBS from cartographic databases or proximity to/availability of managed public spaces such as urban parks. Many of these studies aggregate GBS, effectively classifying GBS as all non-built environments, and draw no conclusions about access or usage.[115] Typologies are similarly varied.[115] Park and open space management often uses a simple four-level hierarchy of parks;[116] effectively, a green-space typology based on the size, catchment area, nature of the resource and the type of facilities.[117] Planning classifications typically focus on the function of GBS. Research studies employ varied typologies, ranging from 2 to 19 types.[118–120] In addition, many existing typologies consider only urban GBS.[117,121] Strategic approaches to planning GBS require more comprehensive approaches that categorise the full GBS resource, urban and rural. Such typologies recognise the importance of a wide range of GBS and facilitate cross-disciplinary understanding.[122]

Objectives

Our objective was to co-produce a UK-relevant typology for urban and rural GBS, primarily for this study but potentially translatable elsewhere. This addresses the need for typologies which facilitate crossdisciplinary and intersectoral work. Our intention was to develop a typology that:

1.   was co-created using expertise from policy and practice
2.   can be operationalised using satellite data and GIS data available in the UK
3.   is relevant to, and can by operationalised by, researchers and policy/practice users.

Methods

We developed the typology using an iterative process (*Figure 2*).

Literature review

We reviewed academic publications, grey literature and policy documentation to collate GBS definitions and typologies. We searched ISI Web of Science databases up to 8 August 2018 using used the same search criteria as a review of green space definitions between 2009 and 2014.[23] We excluded conference proceedings, reports and planning policy documents. We extracted GBS definitions used by UK studies to align with our study setting and noted the geographic location of the study. Where we could not identify an overt definition, we included any examples used in the papers. In addition, we reviewed the main policy literature used in Wales, including the Fields in Trust guidelines, Recreation Opportunity Spectrum, Future Landscapes and WHO.

Stakeholder workshop

In October 2018, we held a semistructured, in-person public involvement workshop for stakeholders that manage GBS and encourage GBS use in Wales. The workshop focused on challenges of existing
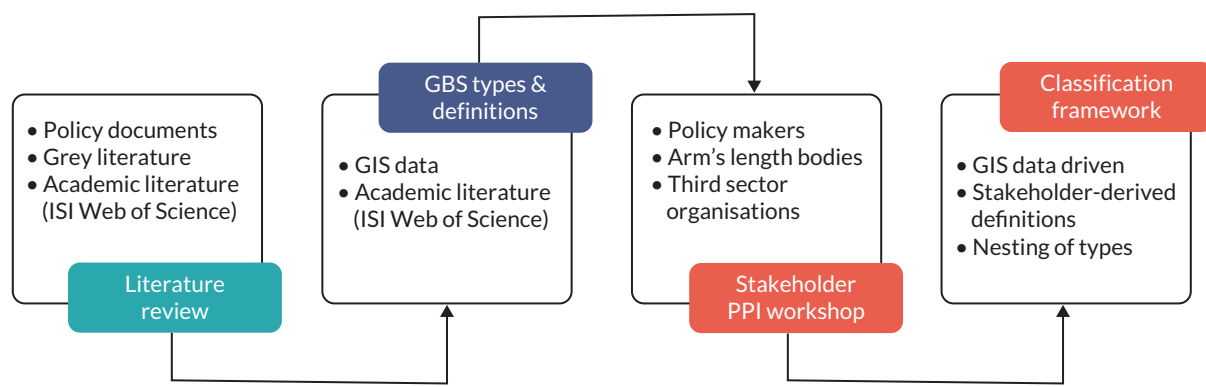
FIGURE 2 Process for developing the typology.

definitions and typologies, potential solutions and implications for working with GBS and data, defining GBS and use of evidence. We invited planning professionals from the 22 Welsh unitary authorities ('local government'), representatives from Welsh Government and arm's length bodies, and individuals involved with third-sector charities. A project summary was shared prior to the event. All participants were offered remuneration in line with INVOLVE guidelines.

Attendees were purposefully allocated to four groups to provide a mix of job roles and expertise. We introduced the project aims and three main discussion topics: (1) challenges and solutions, (2) defining GBSs, and (3) how evidence is used in policy development. Each discussion was designed to elicit feedback that could be used to develop a GBS typology meeting our objectives. Two project team members joined each group, one to chair and one to document the discussions in writing. Those unable to attend the workshop were invited to provide feedback via email.

Results

Literature review

We identified eight articles published between January 2014 and August 2018 to add to the 125 papers identified by Taylor *et al.* for the period 2009 to 2014.[23] We found that policy literature included little information on different GBS types. Definitions of GBS incorporated one or more of availability, accessibility (legal access and perceptions of accessibility), usage or characteristics of a space. Natural England defines accessible green space as 'places that are available for the public to use free of charge' and natural green space as 'places where human control and activities are not intensive'.[123] Definitions differ in what is counted as green space, with certain types of landcover, domestic gardens for example, included by some definitions but not others.[23] This is often related to (private) ownership. NRW defines green space as 'shorthand for both vegetated land – parks, community gardens, trees, woodlands and hedges, informal spaces, allotments and food growing sites, areas of water, such as rivers, canals, lakes and ponds and the seashore'.[22] Blue space is a more recent concept, often defined as visible water bodies 'rivers, lakes and the sea'. Blue space is usually considered a subtype of green spaces, with green spaces potentially including blue features.[114]

We reviewed national planning/policy typologies used in Wales or the UK. Many documents did not define GBS. Where definitions were made, national-scale GBS typologies often classified GBS into broad habitats (e.g. semi-natural grasslands) based on landcover;[124,125] urban areas were considered one landcover type with no further breakdown. Other typologies used limited categories based predominantly on land use/function, such as recreational, amenity and public open spaces.[126,127] Rarely, typologies incorporated the intended age range of users (e.g. provision for children and young people).[120] Many typologies focused on urban GBS.[117,121]

Several policies and reports promote the value and use of GBS in Wales or the UK, typically providing recommendations or standards for availability and accessibility.[128,129] One example is Natural England's Accessible Natural Greenspace Standards, which provides a four-level typology based on distance/size and also considers the quality of the spaces.[130] Fields in Trust has a long-standing recommendation of 6 acres (2.4 hectares) of accessible green space per 1000 head of population.[131] Several typologies originating from the policy and planning sector also consider the availability of different forms of GBS. NRW has sought to implement principles from the Recreation Opportunity Spectrum, which offers a framework to support GBS managers to provide diverse GBS settings for public recreation.[132] Policy guidelines that are widely applied to promote the availability of GBS for health and well-being are often based on arbitrary distance/size cut-offs (e.g. the percentage of the population living within 300 m of a green space of at least 1 hectare).[123,131,133]

Public involvement workshop

Twenty-seven professionals from local government planning departments and third-sector charities, and two individuals from local park engagement groups participated in the workshop. Participants represented 12 different unitary authority planning departments, Welsh Government, the Office for Future Generations, NRW, Sport Wales and Public Health Wales. The third-sector charities Keep Wales Tidy and the Canals and Rivers Trust were represented. Five members of North Wales unitary authorities participated via email. All 22 unitary authorities were engaged via the workshop or email.

Many participants reported that there is no consistency in definition or data sources between agencies. Participants from third-sector charities highlighted that variation in GBS definitions and not knowing which spaces relate to which health outcomes were barriers to advocating for and implementing change. Lack of information related predominantly to social indicators such as safety and usage, rather than to the specific type of GBS. Such factors are rarely integrated into typologies or classifications, nor are relevant data available at scale. However, participants felt that it was important to think about how people perceive places and to ask questions such as: 'Is a park considered to be safe?' or 'Is the space maintained?' It was recommended that the co-created typology classifications should provide a framework to gather information on peoples' perceptions of specific spaces or types of spaces.

Workshop participants noted the need for nuanced consideration of available GBS using more comprehensive typologies. Currently, using available data and applying relevant – if crude – typologies, it can be possible report achievement of guidelines at a city-scale on availability of GBS (e.g. whether they conform to the common standards from Fields in Trust, Accessible Natural Greenspace Standards or WHO). However, they noted that using GBS definitions which overlook finer-scale inequalities can lead to inaction because the guidelines appear to have been met.

Participants considered that differentiating between types of GBS was important for evaluating how different spaces influence health and behaviour. They suggested that the typology should include more diverse forms, or understandings, of GBS such as green corridors. Multifunctionality was also considered important, and the limitations of existing use-related typologies were noted. Further, participants highlighted that different spaces have value for different demographics of the population. For example, one participant suggested that intended age ranges for spaces could be included in the framework. Participants extensively discussed quality as an important element for the framework, but acknowledged that quantifying quality is complex and, in many cases, subjective, individual and context dependent. One suggested solution was that the typology could identify spaces that are nationally recognised (as high quality, e.g. green flags). Finally, several participants raised whether and how the revised typology could capture GBS change over time.

A point of significant discussion was how to manage differences between urban and rural areas. Participants reported that GBS typically differs between urban and rural settings, as does availability and accessibility. They felt that 'countryside' as one 'type' was too broad and that this should be expanded to consider the diverse types of countryside activity, and the specific needs and usage by people who live in rural areas.

Typology synthesis

The results of the literature review and workshop were combined to develop a new typology with a hierarchical structure enabling different levels of aggregation. The revised typology has three nested tiers (*Appendix 2*, *Table 23*). Tier 1 divides GBS into amenities, functional spaces, seminatural habitats, enclosed and linear spaces; tier 2 provides greater detail on the type of space (e.g. identifying a recreation space); tier 3 contains even greater detail (e.g. identifying play areas within a recreational space). We included GBS transport corridors as a specific category based on workshop feedback. The classification has been designed so that additional planning designations (e.g. green corridors and wedges)[134] can be overlayed to the GBS categories. 'Quality' (e.g. biodiversity, upkeep) and 'qualities' (presence of infrastructure such as toilets or natural features such as lakes) relate to each type of GBS, so we did not add green flags as a category. We recommend that the inventory of GBS resulting from applying the typology be combined with information on characteristics, quality, qualities and environmental value where data are available. The detailed methods used to derive GBS exposure metrics from these data are described in *Chapter 4*.

Discussion

We included perspectives of planning and policy professionals to produce a detailed, nested GBS typology applicable to rural and urban areas. Its hierarchical nature allows types of GBS to be aggregated or broken down depending on the purpose or the detail required. This typology can be applied to a wide range of environment data sources (including open source) to classify GBS. It is suitable for applications including research on the health effects of different GBS for specific groups, and analysis of the spatial distribution of different forms of GBS.[3,84]

Much research classifies green space based on land type or the original data classifications. Using original data classifications can present challenges as these data were not collected for public health research, nor with evidence translation in mind. Many epidemiological studies do not include blue space or distinguish between coastal and freshwaters. In the typology we used as a foundation,[117] blue space was only included as 'open/running water'.[41] Our new GBS typology includes the function of the space, details types of inland and coastal blue space, and is designed to facilitate translation of evidence to policy, practice and action.

Green and blue space can benefit health in many ways, not all requiring direct public access. In line with this, in our framework, GBS may be publicly or privately owned or managed and may or may not be publicly accessible. Land which is clearly private or restricted access (e.g. gardens, farmland) was excluded. Accessibility here reflects potential to access a parcel of land that is included in our typology. We acknowledge that 'accessibility' is in fact a much more complex construct dependent on multiple characteristics of spaces, individuals, communities and transport/pedestrian networks. Our classification is necessarily pragmatic and restricted to the data available at a national scale. However, it allows a nuanced understanding of GBS which can inform the protection, improvement, management, and funding of GBS that are suitably located to promote equitable access and potentially also encourage physical activity. We demonstrate its use in public mental health research in *Chapters 5*, *6* and *7*.

Strengths and limitations

Green space is a predominantly urban definition that is more challenging to apply to rural areas. This is partly because we tend to conceive green space as discrete units of space. In rural areas, linear features such as rights of way are more common. We include rural green space but exclude enclosed farmland due to the homogeneity of land classifications in rural areas and a lack of information on accessibility via rights of way. This typology is grounded in published literature and the perspectives of key stakeholders from across Wales. Although this typology is widely applicable, nuances of the Welsh policy and practice landscape should be considered when implementing it elsewhere. The typology is designed to be operationalised using GIS data available at a national scale to fit the requirements of this project.

This led to some pragmatic choices. For example, not all GBS have a defined access point(s) and there is a paucity of data on GBS qualities, so these were not included. Where data on these qualities and characteristics are available, users of the framework can overlay them for the different GBS types.[135] We included all GBS which met the criteria for each category, regardless of size, as the outcomes for the overarching project were not dependent on any specific use of a space, such as for physical activity.

Conclusion

By incorporating stakeholder perspectives, we created a nuanced, nested GBS typology applicable to a wide range of environment data sources to categorise urban and rural GBS. We use this to classify GBS for Wales (*Chapter 4*). It will be particularly useful for research to understand how different groups may be affected by, and benefit from, distinct types of GBS and which are most beneficial to their health and well-being.

Chapter 4 Creating exposure and access measurements to green and blue spaces for every household in Wales

Introduction

A key objective of our study was to calculate longitudinal GBS exposure variables for every household in Wales from 2008 to 2019. There is no universal method for defining the characteristics of GBS that are most important for our mental health. Consequently, research studies have heterogeneity in measures of GBS exposure, with numerous methodological approaches for quantifying exposure to 'greenness'; for example, using satellite imagery to measure regional NDVI;[136–138] self-reported exposure;[37,139,140] and the use of land-cover maps.[39,141]

In this chapter, we discuss how we used available map and other available geographic information to calculate longitudinal GBS exposure variables for our study area (*Figure 3*, *Appendix 3*, *Table 24*). We also describe our findings on national GBS exposure for the study period. Our methods were influenced by our stakeholder engagement; we included characteristics important to stakeholders where possible (i.e. characteristics were available in the routinely collected environmental data).

Measuring ambient exposure to green and blue spaces

We undertook extensive work to define one primary and several secondary exposures that could enable us to answer our primary research questions. This included a literature review to evaluate how previous research studies measured exposure to GBS and scoping of datasets to decide on an appropriate temporal scale. We define our exposures as:

- Primary exposure: annual mean EVI averaged using a 300 m buffer centred on each household location
- Secondary exposures: count of GBS (stratified by type), nearest GBS and average distance to GBS, all within 1600 m of each home within 1600 m of each household location using a network model.

Data source selection

Our primary exposure is defined as temporal changes in ambient green space within 300 m of each household in Wales. Multiple data sources (e.g. aerial photography, land use data and satellite imagery) capture greenspace, but very few have the spatiotemporal coverage required for this study. Satellite imagery is suitable in terms of spatiotemporal coverage and availability of data which are free at the point of use. We exploited the rich satellite imagery data to derive longitudinal measures of greenness as our primary outcome. These satellite-derived measures have long-established methodologies and have been widely applied in research studies.[142]

We chose to use Landsat satellite data (*Appendix 3*, *Table 25*) since they offered the best compromise between cloud-free images, appropriate imagery resolution and temporal coverage (for detailed rationale, see *Appendix 3*, *Figures 30–40*) downloaded satellite data from Landsat projects 4, 5, 7 and 8 using the US Geological Survey EarthExplorer online tool.[143] Owing to the longitudinal design of this study, we required satellite data from 2008 to 2019. Our project straddled different operational lifetimes of remote sensing platforms meaning that no single open-data remote sensing platform had continuous coverage across our study period. To overcome this issue, we downloaded and created composite images from the Landsat project.
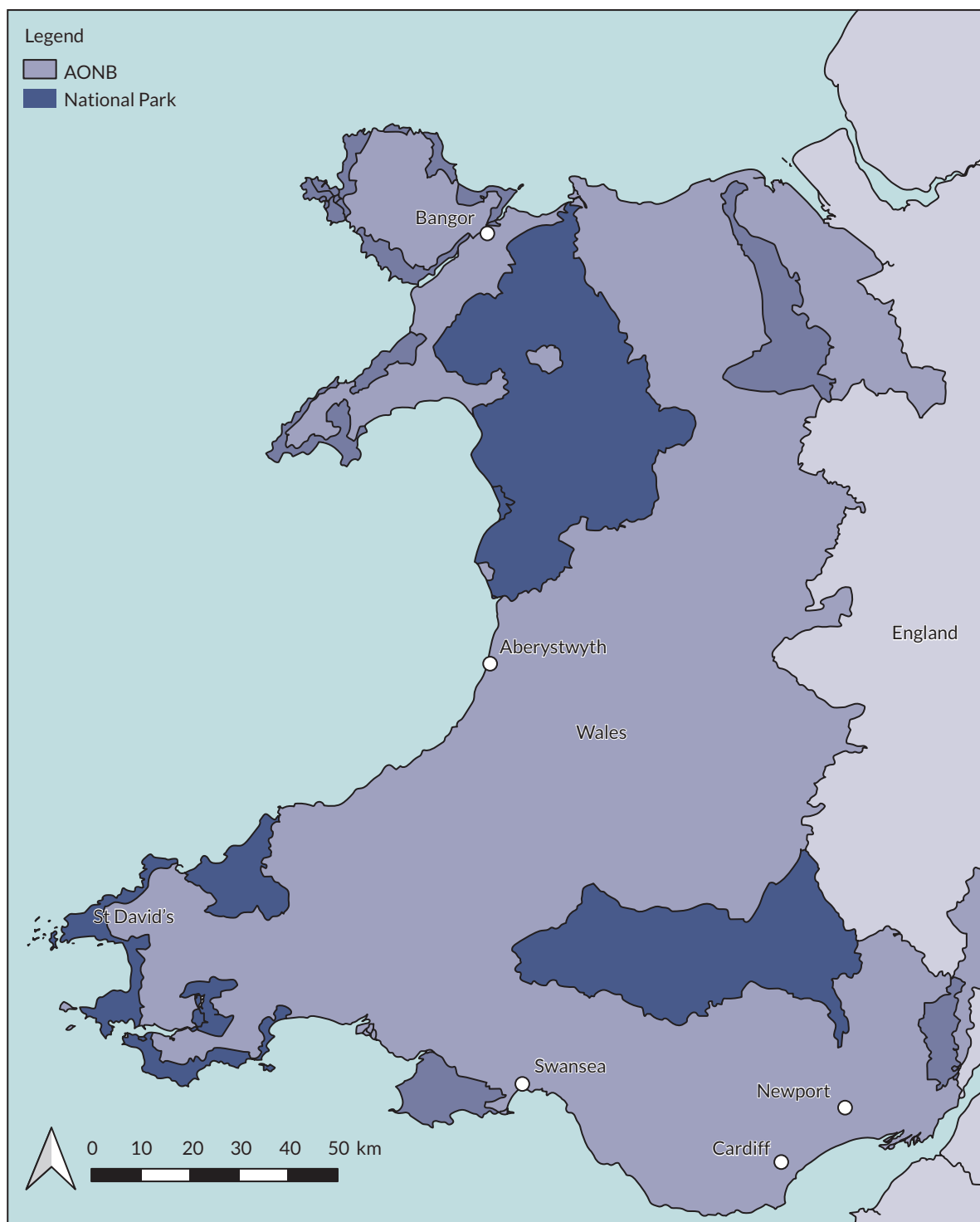
FIGURE 3 The study area including locations of the five main cities, national parks and areas of outstanding natural beauty.

Exposure metric selection

Using these data, we created two measures of greenness density – EVI and NDVI for every year during the study period. We selected EVI as the primary exposure outcome instead of NDVI because EVI reduces both atmospheric and soil background noise. NDVI is sensitive to topographic and meteorological effects; this sensitivity was an important consideration for this study because we do not have satellite imagery for the same time of day and year. *Appendix 4* (*Figures 41–42*, *Tables 30–33*) documents the sensitivity analysis between the NDVI and EVI measures that informed our decision to define the primary exposure as mean EVI within a 300 m buffer of the home location.

Methods

We used Landsat platform versions 5–8 to create a yearly measure of EVI for every household in Wales (*Figure 1*). We selected images from springtime to minimise the number of poor-quality images due to cloud cover. Cloud-free images taken between May and July were downloaded for each year. We preprocessed the images using the Semi-Automatic Classification Plugin tool in QGIS.[144] We applied DOS1 (Dark Object Subtraction 1) atmospheric correction to each image as recommended by Young *et al.*,[145] and calculated EVI for each image using the vegetation index GRASS (Geographic Resources Analysis Support System) tool in QGIS.[146] We also created cloud masks for each image using the Cloud Masking for Landsat Products plugin.[147] We used the cloud masks to set pixels covered by cloud in the satellite imagery to NULL to prevent these values from influencing the final greenness density. We produced annual composite images of Wales in QGIS by mosaicking EVI images together for the same year (see *Appendix 3*).

For each household location in Wales, we created a linear (as the crow flies; Euclidean) buffer of 300 m. For coastal households the buffer was clipped to the coastline to avoid underestimates of green space (blue spaces were captured as described later). Using this buffer, we performed an intersection analysis with the EVI calculation to estimate the density of vegetation at yearly intervals during the study period. EVI theoretically records landcover as values ranging from −1 to +1, with healthy vegetation values found in the 0.2–0.8 range. Values below 0 represent water bodies and were generally excluded by clipping coastal areas.

Geographic information system modelling to measure access to green and blue spaces

Data sources

We used vector data from multiple sources to calculate access to GBS (*Appendix 3*, *Table 26*). Sources included data from OS, local government audits and open data sources. The following section describes the sources of data.

MasterMap

We used OSMM Topography Layer,[148] the most detailed, continuous view of the UK landscape, which has mapped UK surface with land parcels for natural and man-made features. We also used the OSMM Greenspace dataset,[149] which is the UKs most comprehensive green-space dataset for England, Wales and Scotland. These data were accessed under the Public Sector Mapping Agreement.

Local government audits

The TAN 16 supplements 'Planning Policy Wales and provides guidance for local planning authorities on recording sport, recreation and open spaces.[100] Open spaces include public parks and gardens, outdoor sports facilities, play areas, allotments and water (e.g. ponds, rivers and reservoirs). Local authorities are legally required to record and manage data on their open spaces through TAN 16, but there is no guidance on what data to specifically record and how to manage the data. We requested TAN 16 data from all 22 Welsh local authorities: 14/22 shared their data. This process is documented in detail in *Appendix 3*.

Open data

We also collected vector data from open-source portals. These included: Lle (Welsh Government and NRW data repository),[150,151] where we used forestry and urban tree polygon data, and OpenStreetMap,[152] from which we extracted park facilities such as kiosks, public toilets and road and footpath networks.

Harmonising geographic information system data

We harmonised the GIS data described above to create a baseline dataset of all potential GBS in Wales. We brought together different spatial data types, reprojecting data to consistent geographic reference systems and spatial scales and classifying data into broad types of GBS. The data were managed and harmonised in PostGIS.

Operationalisation of the green and blue spaces typology

The GBS typology is a theoretical framework to categorise land parcels in Wales into GBS which people can potentially access, both visually and physically, to understand the impacts and benefits to mental health and well-being. The typology does not directly map to predefined GIS data or categorisation found within these data. To operationalise the typology, we applied a set of criteria-based processes and pragmatic choices.

First, we applied a set of criteria-based rules to rationalise each of the land parcels in the GBS data using the typology (*Figure 4*). To avoid duplication when combining multiple data sources, only vectors that did not overlap with the OS dataset were added from TAN 16 local authority data and the open data sources. We defined both GBS as land parcels to maintain a consistent definition of access to GBS.

We next applied the typology to the data using a three-tiered approach, based on feedback from the stakeholder group, with tier 1 being the least detailed and tier 3 the most. *Figures 5* and *6* show the differences between tier 1 and tier 3 for the area around Swansea University's Singleton campus. SQL code for operationalising the GBS typology is available at https://gitlab.com/envhe/BGS.git.

Once each land parcel which met the typology criteria was categorised into the typology subgroups, we created potential access points to measure distances from each home to all GBS. Most GBS captured in our typology are not managed and do not have specific access points. We developed a method which accounted for nested GBS and made it feasible to calculate access to GBS distances and counts. Owing to the variety of types (e.g. managed and unmanaged) and morphologies (e.g. nesting of GBS within other GBS) of GBS within the dataset intersections with road or footpath features with a GBS did not always produce an access point (e.g. a linear beach feature with a
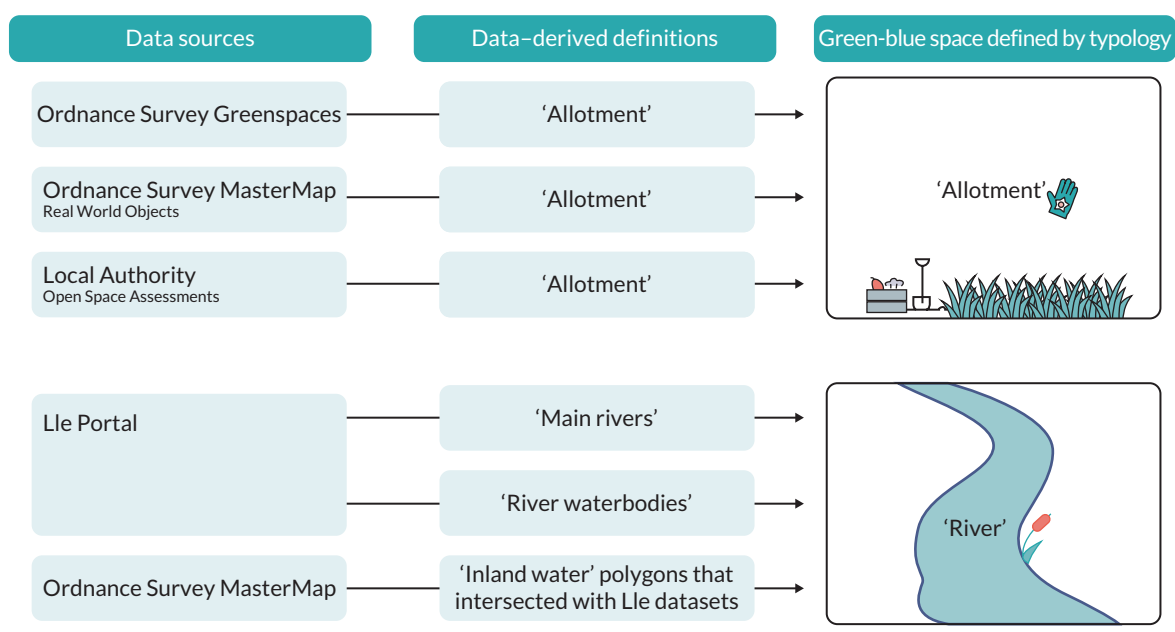
FIGURE 4 Conceptual diagram of applying typology to longitudinal dataset using a criteria-based approach.

FIGURE 5　Tier 1 typology for the area around Swansea University's Singleton campus.

FIGURE 6　Tier 3 typology for the area around Swansea University's Singleton campus.

footpath running parallel). We developed a method for defining access points based on feature types which would act as proxy for accessing each green space. The following criteria needed to be met:

- For each regularly shaped green space a proxy access point for each side of access needed to be defined (e.g. north, south, east, and west).
- For linear features (e.g. rivers, canal, coastlines) an access point will be defined at set intervals to reflect multiple points of access along the length of the space.
- Each access point needs to be snapped to a footpath or road network node to define a point along the network at which the GBS can be accessed.

Figure 7 shows the result of implementing the access point criteria on a sample area of the GBS dataset. It shows how far each access point was moved in metres; this snapping distance was included in the final distance calculations. The resultant data were a set of point locations snapped to the road and footpath network representing potential access points to all GBS in Wales.

Measures of the accessibility of green and blue spaces

Accessibility of GBS is nuanced, involving a combination of proximity, physical environmental characteristics, facilities (such as public toilets) and individual and social/cultural characteristics that interact to impact how accessible and appealing any space is for an individual or community.[153,154] While acknowledging this complexity, this study conceptualises accessibility primarily in terms of proximity to home. We defined two measures of access as secondary exposures; counts of GBS (stratified by type), nearest GBS, and average distance to GBS, all within 1600 m of each home in Wales. Many studies have used only nearest green space as a metric. It is evident from the Monitor of Engagement with the Natural Environment data[155,156] that 1600 m (approximately 1 mile) is the point where a rapid decline in greenspace use is reported.
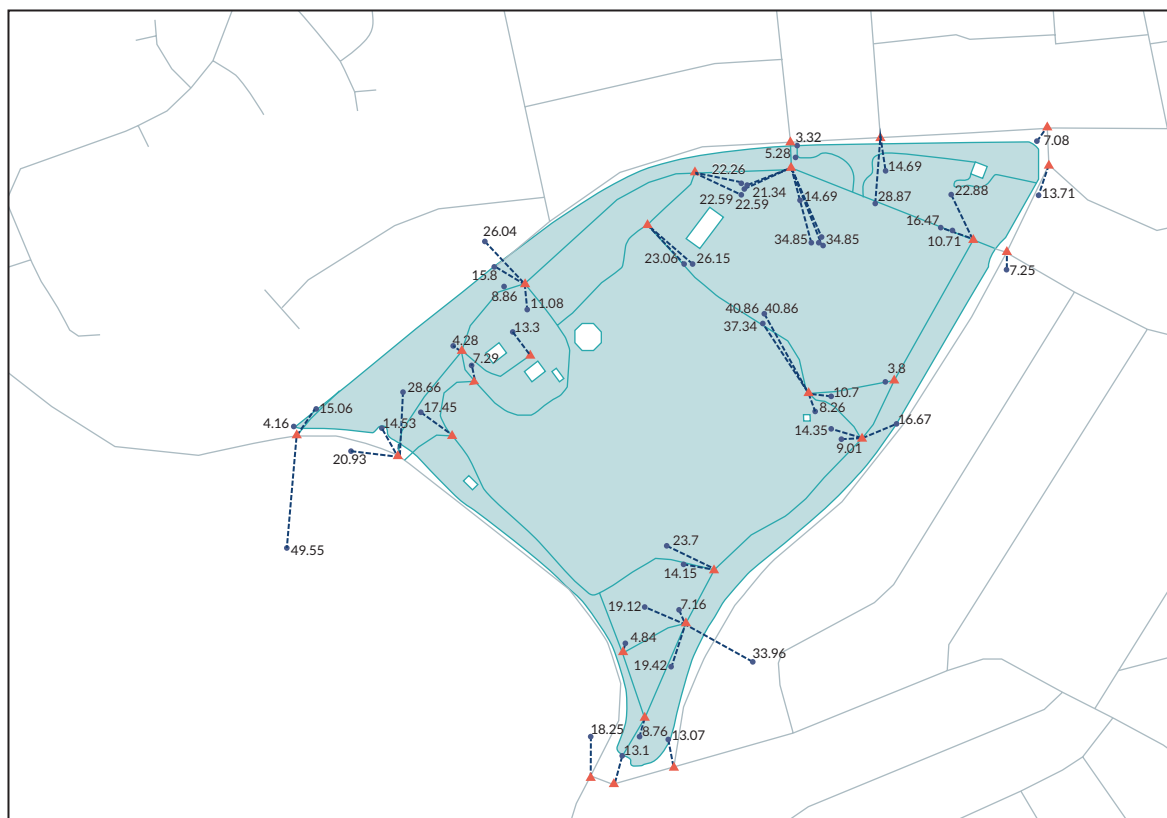
FIGURE 7 GBS proxy access points snapping including distance moved in metres.
circle = original access point, triangle = snapped location – figure distance moved in metres from local road network = blue lines, to an access point within the GBS = green shaded area.

We also captured a variety of other metrics, including average size and nearest GBS size, to explore nuances in accessibility which have been used in our sensitivity analyses (metadata for the secondary exposure calculations are given in *Appendix 3*).

Methods

We created an origin–destination (OD) network (roads and footpaths) distance matrix for each household in Wales (*n* = 1,498,120 households) to estimate the distance to all GBS access points (*n* = 3,242,803) potentially accessible within 1600 m. This resulted in an OD matrix which paired each GBS with each household in Wales to create unique combinations of distances for 772,557,144 pairs of homes and GBS. We de-duplicated each OD pairing so that each household only accessed the nearest access point of an individual GBS. The final distance calculation was a sum of: (1) the snapped distance from household to nearest network node (d_i); and (2) the total network distance (d_j) between d_i and the network location of the snapped GBS access point (d_k) for all GBS within 1600 m of each home.

For use with the health data, we created three GBS access measures: summed total counts, nearest GBS, and average distance to a GBS. We also grouped each OD matrix per household to count by type, resulting in a secondary GBS access measure; a count of all GBS types within 1600 m (*Equation 1*).

$$secondary\ exposure = \sum_{d_i+d_j+d_k \geq 0}^{d_i+d_j+d_k < 1601} GBStype^{1\ldots n}$$

(1)

Equation 1: Access to green and blue space calculation where d_i = snapped household distance, d_j = network distance, d_k = snapped GBS access point distance.

Typological classification of GIS data and calculation of proxy access points were performed in PostGIS using SQL. Subsequent summary counts were calculated in Python 3.6 using a high memory environment (128GB RAM with 8 cores). We managed the data in a secure, dedicated GIS repository to ensure best practice with regards to household-level data and because we produced very large volumes of longitudinal, spatial data for linkage with individual-level health data. The final data were anonymised through our trusted third party using a split-file process.[157,158] Identifiable details were removed from the dataset and replaced with an ALF or a RALF.[159] An ALF is assigned if the data are individual-level records or a RALF assigned if the data are address-level data. We linked the GBS access measures with our study cohort (*Chapter 3*) where we also describe the GBS exposures by population characteristics (e.g. deprivation, age, sex).

Results

We report results for our primary exposure (annual ambient greenness measured by EVI) and secondary exposures (access to GBS within 1600 m). We found that most households in Wales had some greenness within their local vicinity and have GBS access (as defined by the typology).

Primary exposure: enhanced vegetation index

The spatial distribution of ambient greenness across Wales is consistent with the theoretical principles of an EVI estimate, with rural areas appearing to have higher ambient exposures. We found some variation over time in the spatial distribution of EVI estimates (*Figure 8*) but spatially they appear to be random effects with no obvious clustering. We captured average EVI scores per household every year during the study period (see *Appendix 3*, *Table 27* for annual average EVI aggregated to LA level).

Rural regions have a greater EVI score every year compared with urban areas (*Figure 9*, *Appendix 3*, *Table 28*); however, there has been increased greening in both settings over the study period. In 2012 there is a noticeable dip in the EVI in rural areas, this corresponds with the lower mean EVIs visible in *Figure 8* for the northwest area of Wales. Cloud cover meant that every satellite image taken in 2012 had poor visibility.
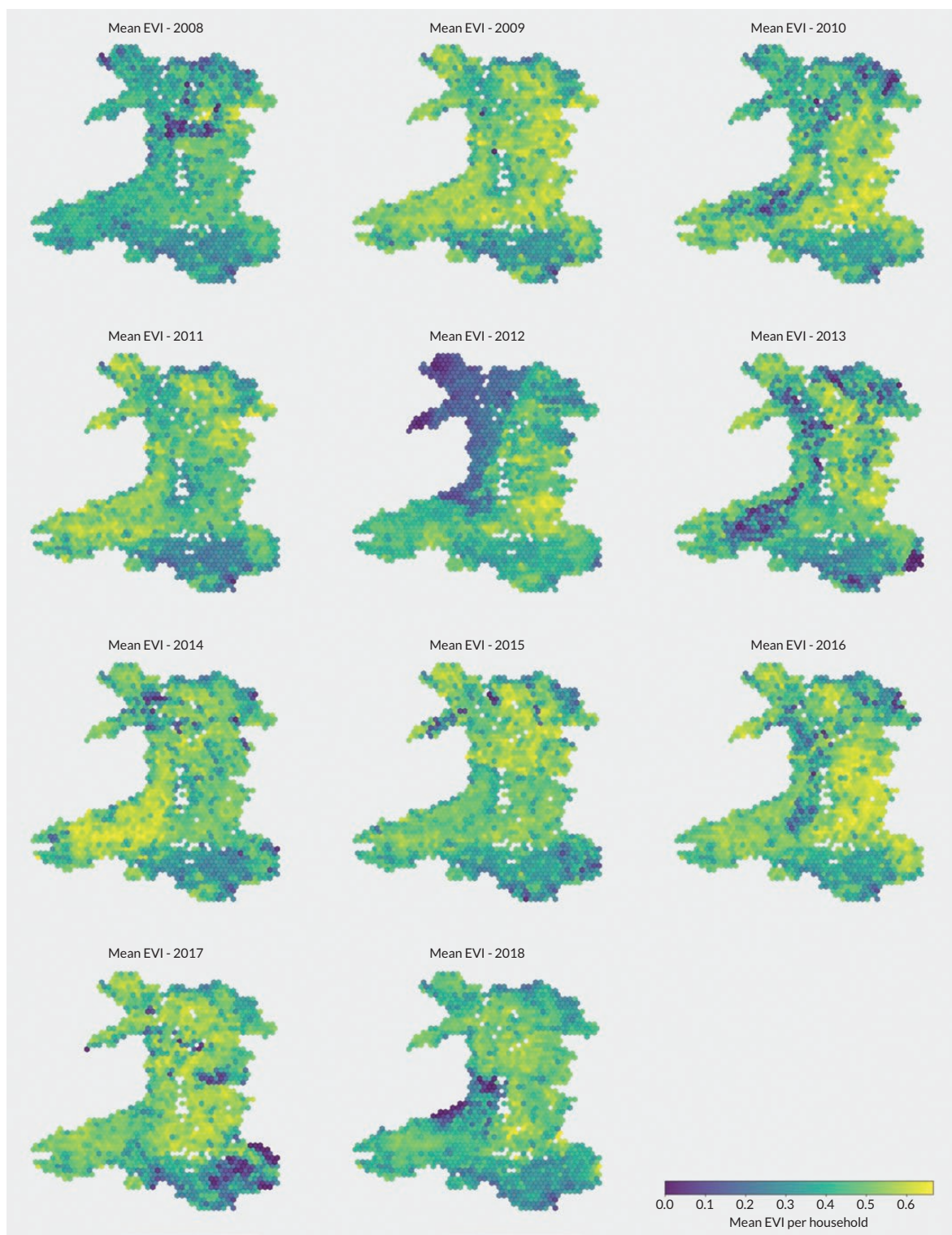
FIGURE 8 Aggregated EVI measures 2008–2018.

Spatial distribution and access to green and blue spaces in Wales

Figures 10 and *11* show the distribution of potential access to GBS by household, aggregated to hexbins to maintain disclosure controls. Areas where there are too few households to aggregate data are shown in white. *Figure 10* shows the distribution of access to green space across Wales with rural areas experiencing access to less formal green spaces (defined by the typology) than those found in coastal and more populated areas. The northwest of Wales has some of the highest access to green space but is
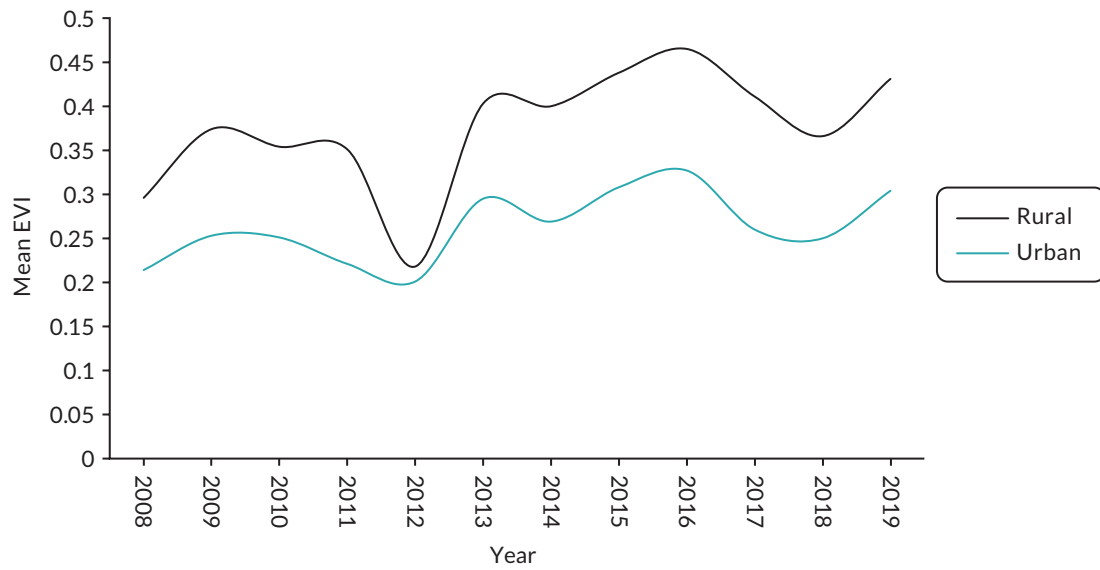
FIGURE 9 Temporal variation in household-level mean EVI by rural/urban location.
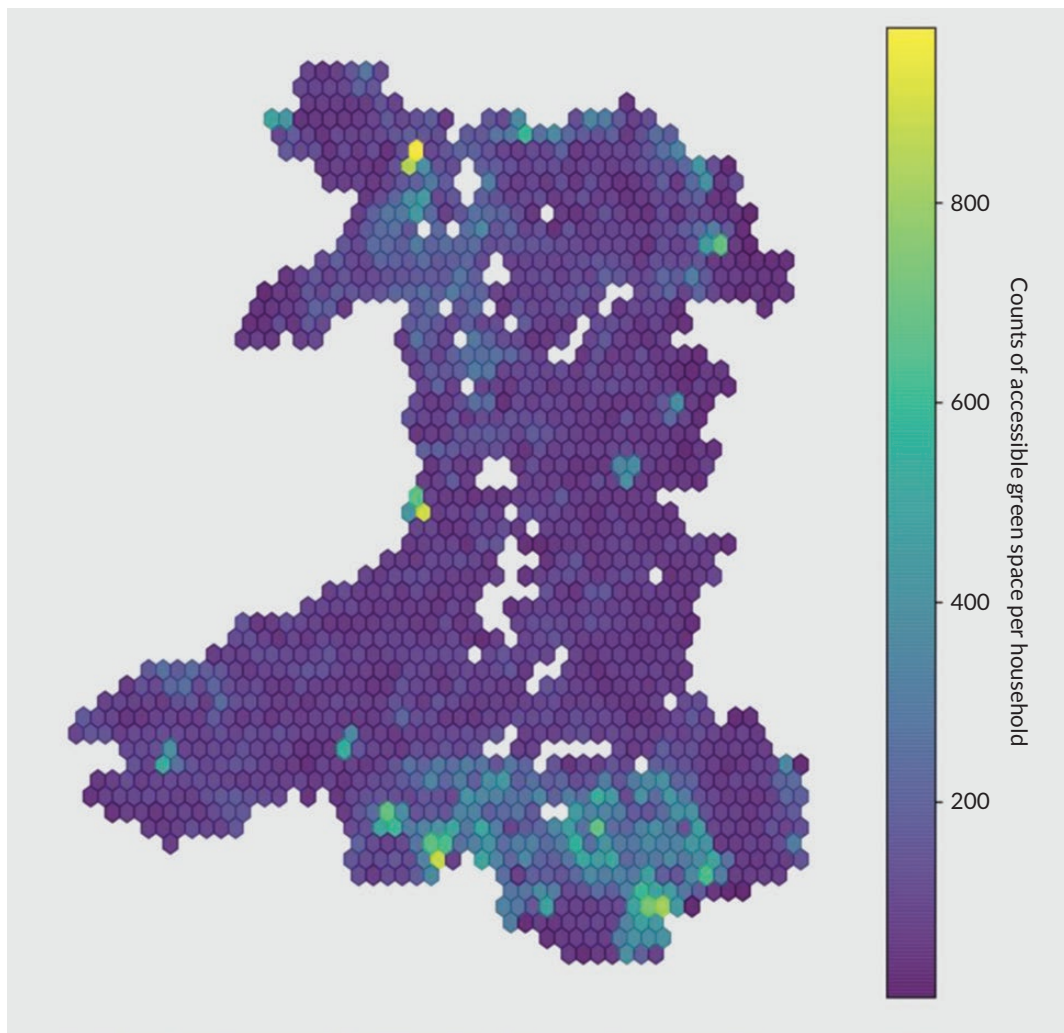
FIGURE 10 Aggregated access measures to green spaces per household within 1600 m.
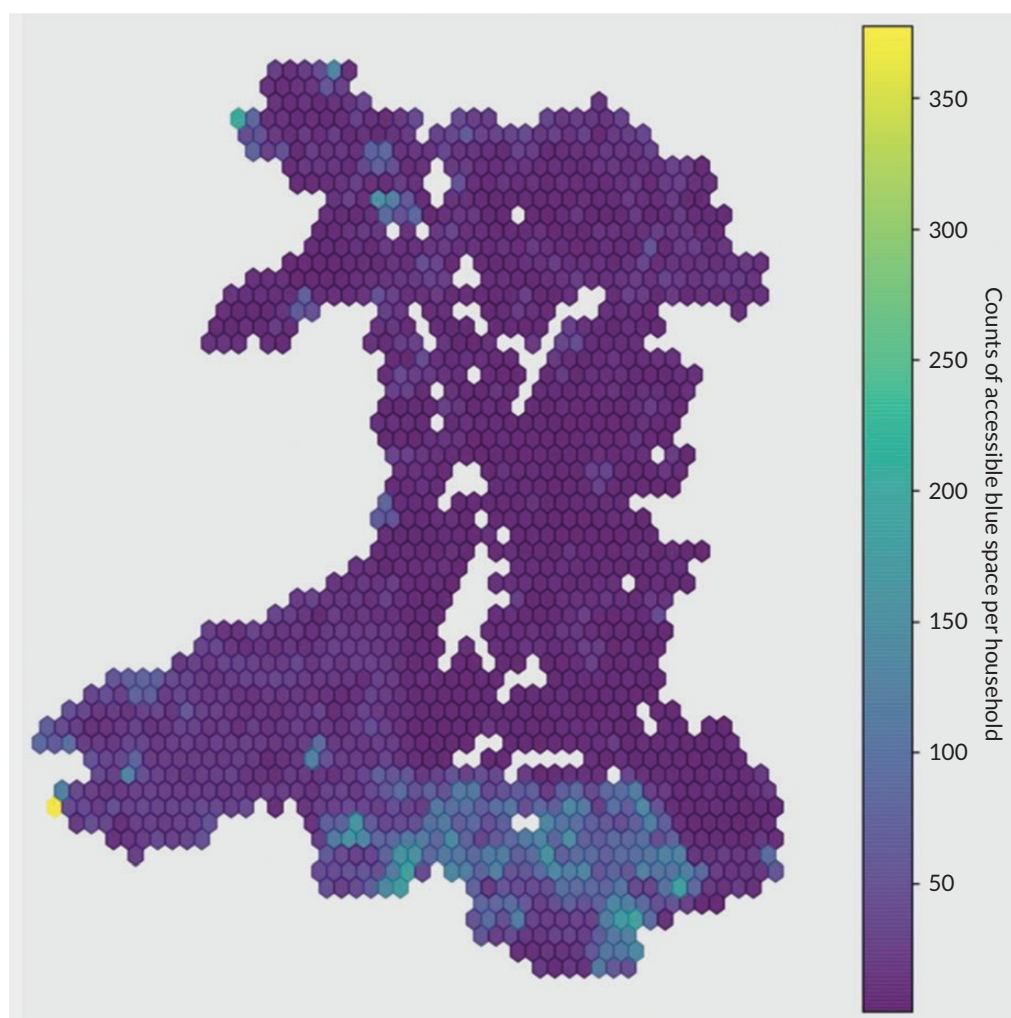
FIGURE 11 Aggregated access measures to blue spaces per household within 1600 m.

also very rural and has formal access to national parks and coastal areas for the population living in the vicinity. Areas which are managed as national parks or sites of special scientific interest (e.g. Snowdonia, Brecon Beacons, Pembrokeshire and the Gower Peninsula; formal conservation designation in the UK, of particular interest to science due to rate fauna, flora or important geological features) are more sparsely populated and have poorer household levels of access. *Appendix 3*, *Table 29* highlights that 97% of land cover does not change in Wales from 2007 to 2015. Increased access to blue space is found in the more populous regions of the South Wales valleys, cities and coastal towns.

Capturing changes in potential access to green and blue spaces

The typology we developed to classify land parcels into different categories of GBS can be applied to the same data sources from different years, where they exist, due to the standardised method for capturing and storing information. To assess the feasibility of detecting change between two epochs of data, we applied the typology to OSMM data to capture greenspace availability in 2012 and 2018. We found that only limited changes in the access profile to GBS at a population level using different epochs of the GBS access data. There do, however, appear to be small regional variations which would impact at a household level (see *Figure 12*).

At an individual household level, changes are more apparent, with some households recording substantive changes in the numbers of GBS available.

Table 4 gives an example of a household which has shown substantive changes between 2012 and 2018.
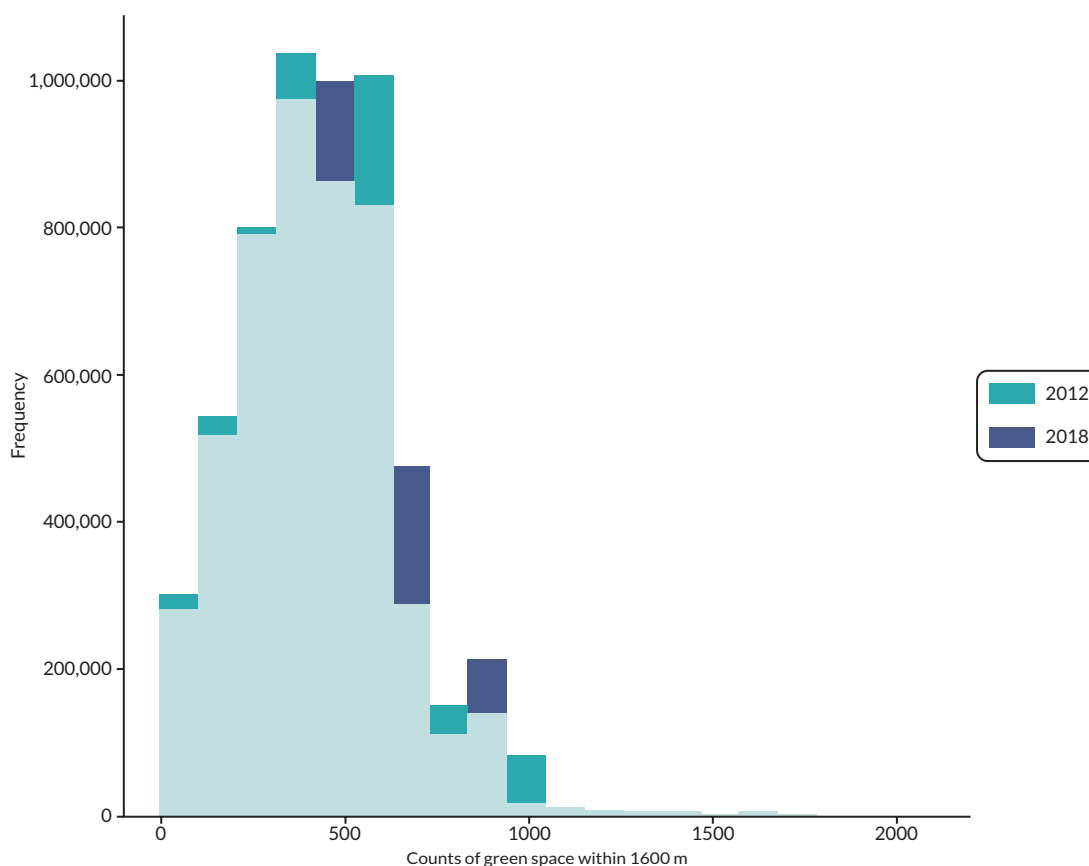
FIGURE 12 Density plot showing distribution of number of green spaces potentially accessible in 2012 and 2018 for ~1.5 million households in Wales.

TABLE 4 2012 and 2018 accessibility results for a single household in the valleys area in Wales

	Year	
Variable	2012	2018
All counts	60	47
Average distance (m)	658.5	701.5
Nearest distance (m)	130	121
Average size (m²)	47,454	66,548
Nearest type	Coniferous	Grassland

Upon exploration of the data, it became apparent that these changes did not necessarily represent changes in the built or natural environment around a home, we discuss this in further detail in the next section.

Discussion

We used a variety of data types and sources to create a national scale environmental dataset to understand the associations between the environment and mental health and well-being. We have shown that satellite data can be used to derive comparable, longitudinal measures of greenness and access to GBS. While computationally intensive, we also showed it is feasible to create indicators of GBS access at a household level for a whole nation, using advanced GIS techniques. There are several limitations linked to the data, which we discuss here.

We found that there are challenges in harmonising (1) where stakeholders identified developments for example including more dimensions in our data as a proxy of quality; (2) the data that are recorded; and (3) the data that are available to researchers. The long temporal and large spatial scales of this study meant that the highest temporal resolution we were able to create comparable exposures for was annual measures, and micro-features of the environment were not included in our exposure variables.

Estimating EVI using satellite imagery was a challenge because most satellite sensors are passive and of relatively low pixel spatial resolution. Wales does not experience many cloud-free days. As such we adopted a flexible approach to estimating EVI. For example, we used different Landsat sensors to enable EVI measures to be calculated throughout the study period. While this is not an unusual approach in the environmental sciences,[160] it limited our ability to capture seasonal changes or to fully adjust images to remove clouds. Invariably, this will have caused slight variations in EVI sensitivity, despite corrections, and may have caused the 2012 anomaly in Northwest Wales. Further, EVI as a measure is inexorably linked to seasonal variation in peak vegetation greenness which, dependent on satellite coverage, atmospheric conditions, and cloud cover, means we cannot be certain whether small changes observed in household levels of ambient greenness were due to seasonal or permanent changes.

Overall, we saw a general 'greening' throughout the study period. However, Welsh households did not see a large spatial variation in mean EVI values. An example of how EVI values varied across an urban and rural area in Swansea is included in *Appendix 3*, for illustrative purposes. Each EVI value represents a unique combination of vegetation reflectance and seasonality. Identical EVI values may represent different greenspace profiles. Rural areas generally had higher EVI values and less access to GBS as defined by the typology. There was more blue space accessible in the more urban South Wales valleys region. This may reflect Wales' industrial heritage, with access to canals and ponds accounting for the differences.

Our secondary measure of GBS access relied on data that were less encumbered by the Welsh weather. However, there were a different set of limitations for these vector data. First, we had to create proxy access points. This was necessary due to the lack of information on which GBS were accessible through an entrance only (i.e. were enclosed by a wall or fence) and which had an open boundary. We were unable to obtain data on rights of way in rural areas and therefore had to exclude privately owned farmland, which may have sections which are publicly accessible to walkers. Combined with the inclusive approach to the typology we consider the access measures to be indicative of potential access to GBS in Wales, and a pragmatic approach to measuring the heterogeneity in GBS across Wales.

Second, while the OSMM Topographic Layer is a very rich dataset, it is not designed to capture longitudinal changes. During our study, we discovered that the continual improvement programme undertaken by OS[161] meant that the database behind the OSMM Topography Layer continually changed. This resulted in changes in land use that were not reflective of real changes. Artefacts were introduced making these data unsuitable for longitudinal studies. The majority of temporal change we saw in our study for green spaces fell into a 'category B' update policy which triggered an update when an area of over 25 Ha experienced a change in land use (e.g. brownfield site to housing estate).

Most of the green space potentially accessible from each home in Wales is less than 25 Ha. For 95% of households the nearest green space is less than 10 Ha. The OS has changed their map update frequency, resulting in many land parcels being updated with an unknown and infrequent revisit cycle of between 2 and 10 years. This has likely created insufficient updates to enable us to create a longitudinal measure of GBS access. We instead used a baseline measure of 2012. In our work, longitudinal variation in GBS access originates only from people who self-selected to move home, while EVI-based exposure measures were updated annually for everyone even if they did not move home.

Classifications of land use within OSMM are also subject to updates which reflect continual improvements in land use classification rather than changed land use. This can be seen in *Figure 11*,

where changes in the categorisation of land use (e.g. grassland to arable land) and changes in the delineation of managed forestry areas resulted in significant differences in the land parcels captured by the typology framework but do not reflect real changes in GBS types available in the local community. Many of the changes we observed in GBS datasets are due to reclassification of land parcels. Examples of change in land classification vary but typically represent improved accuracy in the representation land use. In *Figure 13*, land parcels have been reclassified from grassland to arable lands. Grassland was included in the typology; arable land was not due to unknown rights of way across much of the arable land in Wales. Fundamentally there has been no change in GBS and access to this land will not have changed, but this would be misrepresented in the GBS access measure.

A further example is where there have been changes to classification of land parcels for no logical reason. *Figure 14* shows a playground which has been classified as such in 2012. However, by 2018, this playground has disappeared from the typology due to a database update at OS, which now classifies this as 'general surface'. The space no longer meets the inclusion criteria for the typology, thus appearing as a loss of greenspace when modelled.
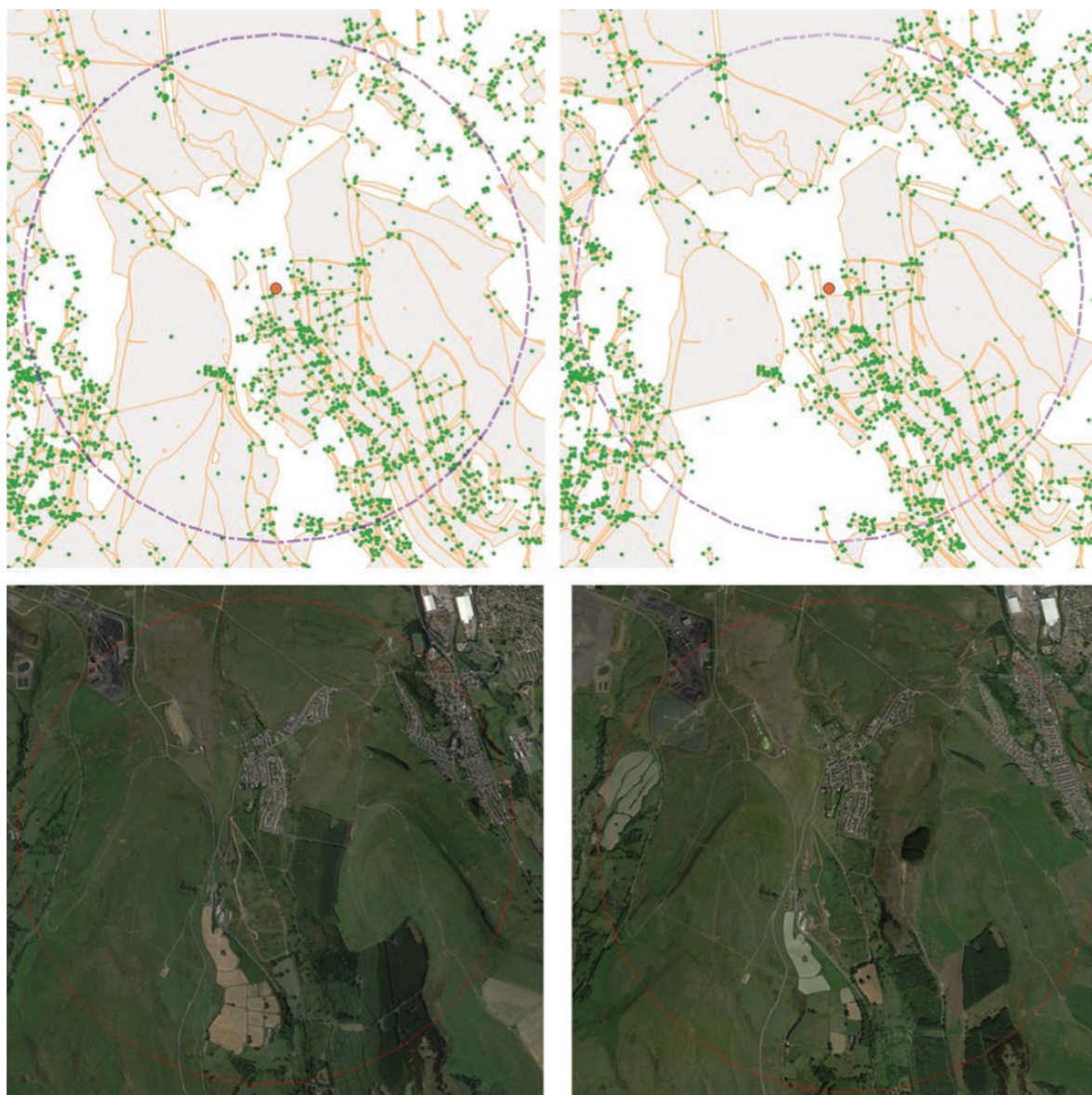
FIGURE 13 Greenspace landcover based on typology and accompanying satellite imagery. Recorded change in OSMM: grassland to arable 2012 (left) and 2018 (right).

FIGURE 14 Greenspace landcover based on typology and accompanying satellite imagery.
Recorded change in OSMM: playground to general surface 2012 (left) and 2018 (right).

Owing to the uncertainties presented by these changes in the core dataset and a lack of other longitudinal data from other data sources, we were unable to confidently model change for our secondary exposure of GBS access. Given that most substantive changes in exposure for the cohort will have arisen through moving home, the impact of this limitation is likely to be relatively small.

Conclusion

In this chapter, we have detailed methods to objectively measure two exposures to GBS: ambient exposure around homes and potentially accessible GBS. These methods are based on the typology detailed in *Chapter 3* and use a range of GIS methods to create environmental exposures at household levels for a whole country over an 11-year period. Uniquely, these data were anonymously linked at household level to routinely collected primary care health data and representative nationally collected survey data. Our primary exposure, annual measures of EVI, were utilised for our main longitudinal analysis. We have reported that vector data do not reliably reflect *change* in GBS over time and therefore, for our secondary exposure, our longitudinal analyses were restricted to allocating 2018 access measures to individuals who move home during the study period.

Chapter 5 Green and blue spaces and common mental health disorders: time-aggregated analysis

Introduction

Despite known limitations, time-aggregated analysis can be a valuable approach for longitudinal studies, avoiding methodological and interpretive challenges of more complex analyses. Here, we apply time-aggregated analysis of individual-level data for our cohort to assess the extent to which cumulative GBS exposure may be associated with preventing a CMD. This approach enabled us to consider GBS exposures prior to CMD outcome. Our work exploits the availability of environmental metrics at the household level, reducing ecological fallacy, occurring when characteristics of areas are incorrectly attributed to individuals. These data were previously inaccessible to national level routinely collected health datasets due to governance restrictions in the UK.

We describe the sociodemographics, GBS exposure and CMD outcome of our cohort. We present model results, followed by a discussion of these results in the wider context, including the strengths and limitations of data and our approach.

Aim and research questions

We aimed to assess the extent to which cohort members' residential GBS exposure and potential access to GBS is associated with preventing a CMD (as determined through use of primary health care services). We did not assess whether access to, or use of GBS, is therapeutic for an existing CMD.

Our research questions are:

Research question 1: Is greater residential exposure and access to GBS associated with the likelihood of seeking help for a CMD in general practice?
Research question 2: Is the association between exposure, or access to, GBS and seeking help for a CMD modified by socioeconomic disadvantage and moving home?

In this work, we focused on spatial–temporal differences in exposure to GBS between different people, at a highly granular level. Quarterly individual-level data on environmental exposures (GBS exposure and access) were aggregated to a single value for the period an individual was (1) in the study, and (2) before being coded as a CMD flag. We explored whether aggregated GBS exposures were associated with the proportion of individuals flagged as a CMD case. CMD cases were identified based on primary care records using a validated algorithm (see John *et al.* 2016 and *Chapter 2*).[103] Hereafter, we describe this population as 'having a CMD', recognising that diagnosed CMD underestimates the population prevalence but is a consistently measurable attribute.

As stated, this chapter focuses on the differences between different people within space, and over time. In *Chapter 6*, we present the results of our longitudinal analyses, using only the differences between the same people through time, moving closer to causality by removing the selective spatial differences between people. The results in this present chapter help to elucidate the relationships between GBS and primary health care utilisation and provide a valuable addition to the field of environmental epidemiology.

Methods

Study design

Our time-aggregated analysis comprised 47 quarters, from January 2008 to October 2019. Only exposure data prior to the first recorded CMD were used. For instance, if an individual first had a CMD in the second quarter of 2013, then data on their GBS metrics up to and including the first quarter of 2013 were aggregated. Ambient greenness (EVI) and GBS exposure were normalised to between 0 and 1 to allow comparability across exposure measures. Results are presented for increments of 0.1 unit.

The time-aggregated cohort analyses used a subset of the longitudinal cohort described in Chapter 2. Individuals from the longitudinal cohort with a CMD in the first quarter in the study were excluded since there were no prior environmental exposure data. All other individuals in the longitudinal cohort were eligible for inclusion in the time-aggregated analysis. Individuals were removed if it was not possible to derive a suitable set of covariates and factors from the longitudinal data due to missing data.

Variables and data aggregation

The primary outcome for this time-aggregated analysis is binary. For each quarter we assigned every individual an indicator for either a high likelihood or low likelihood of having a CMD.

We defined the primary exposure as ambient exposure to greenness, the average EVI value of a 300 m buffer around the household location. Our secondary exposures describe potential access to GBS: number of GBS within 1600 m of the home location; distance to nearest GBS within 1600 m; and average distance to GBS within 1600 m. Individual-level GBS exposure and access summaries were modelled quarterly.

Data from quarters up to and including that before a recorded CMD, or until the individual moved out of Wales, were used to derive a normalised single time-aggregated 'average exposure' value for each individual. All covariates were derived from data in the same time interval (Appendix 4, Table 30).

Statistical analysis

We applied binary logistic regression analyses to assess the associations between exposure and access to GBS with our binary CMD outcome. We used univariate models to assess whether associations were observable between our exposures, CMD and confounders. Where we detected associations, we included these variables in multivariate models. We also stratified the analyses by deprivation, history of CMD, rurality and migration (moving home) to assess effect modification.

In addition to reporting odds ratios (ORs), we also interpret model output by calculating predicted CMD probabilities. This illustrates how the probability of a CMD varies with exposure metrics, and how probabilities differ for men and women of the same age living in the same deprivation quintile.

Sensitivity analyses were performed on the additional exposure and access measures, including models where both the primary exposure measure and access measures were included.

Results

Descriptive statistics

Ninety-one per cent of the longitudinal cohort (n = 2,801,482) were eligible for inclusion in the time-aggregated cohort. Following exclusion of 247,588 individuals due to missing data (Figure 1); the final time-aggregated cohort comprised 2,553,894 people; 10 or 11 years' worth of data were available for almost half (45.4%) of the cohort (Table 5).

Of the 2,553,894 individuals in the time-aggregated cohort, 21.9% ($n = 513,239$) had a CMD recorded at least once. A higher proportion of females had a CMD compared with males (26.8% and 17.2%, respectively; *Table 6*). More than one-quarter of those in the 22–50 years age groups recorded a CMD ('cohort with CMD', *Table 6*). This was approximately twice as high as the proportion in the oldest age group (80+ years: 13.3% with a CMD, *Table 6*).

TABLE 5 Time-aggregated cohort CMD profile – number of quarters with individuals' data in the time-aggregated analyses

	All		CMD		No CMD	
Number of quarters	**(%)**	**(n)**	**(%)**	**(n)**	**(%)**	**(n)**
1–8	14.9	348,862	29.0	148,764	10.9	200,098
9–16	13.1	307,498	22.8	117,230	10.4	190,268
17–24	10.3	239,937	17.6	90,278	8.1	149,659
25–32	8.8	205,920	14.0	71,788	7.3	134,132
33–40	7.5	176,117	10.8	55,400	6.6	120,717
40–47	45.4	1,063,257	5.8	29,779	56.5	1,033,478

TABLE 6 Sociodemographics of the time-aggregated cohort

	All		CMD		No CMD	
Group	**(n)**	**(%)[a]**	**(n)**	**(%)[b]**	**(n)**	**(%)[b]**
Age group (years)						
16–21	363,314	15.5	70,282	19.3	293,032	80.7
22–30	385,643	16.5	102,992	26.7	282,651	73.3
31–40	327,874	14.0	87,550	26.7	240,324	73.3
41–50	338,303	14.4	88,904	26.3	249,399	73.7
51–60	302,774	12.9	67,545	22.3	235,229	77.7
61–70	284,096	12.1	47,105	16.6	236,991	83.4
71–80	201,279	8.6	30,401	15.1	170,878	84.9
80+	138,308	5.9	18,460	13.3	119,848	86.7
Sex						
Male	1,193,240	51.0	205,546	17.2	987,694	82.8
Female	1,148,351	49.0	307,693	26.8	840,658	73.2
Birth in household						
No	2,076,801	88.7	460,262	22.2	1,616,539	77.8
Yes	264,790	11.3	52,977	20.0	211,813	80.0
Death in household						
No	2,151,488	91.9	489,243	22.7	1,662,245	77.3
Yes	190,103	8.1	23,996	12.6	166,107	87.4

Table 6 Sociodemographics of the time-aggregated cohort *(continued)*

Group	All (n)	(%)ᵃ	CMD (n)	(%)ᵇ	No CMD (n)	(%)ᵇ
WIMD						
1 – Most deprived	473,404	20.2	131,843	27.8	341,561	72.2
2	462,492	19.8	110,898	24.0	351,594	76.0
3	480,278	20.5	98,673	20.5	381,605	79.5
4	445,002	19.0	83,768	18.8	361,234	81.2
5 – Least deprived	480,415	20.5	88,057	18.3	392,358	81.7
Moved home						
Does not move	1,439,204	61.5	343,734	23.9	1,095,470	76.1
Moved once	520,509	22.2	99,694	19.2	420,815	80.8
Moved more than once	381,878	16.3	69,811	18.3	312,067	81.7
Rurality						
Urban	1,611,516	68.8	365,931	22.7	1,245,585	77.3
Rural	354,961	15.2	62,285	17.5	292,676	82.5
Town and fringe	375,114	16.0	85,023	22.7	290,091	77.3

a Percentage of the whole cohort in this category.
b Percentage of individuals in this category who sought/did not seek care for a CMD in general practice.

Common mental health disorders and variation in exposure

Individuals with a CMD had, on average, a slightly lower ambient greenness exposure than those who were not identified with a CMD (EVI: 0.27 and 0.29, respectively; *Figure 15* and *Table 6*). The average count of GBS spaces was also greater for people without a CMD (counts: 430 and 445, respectively; *Figure 16* and *Table 7*).

There were differences in the extent of primary and secondary exposure with age. Younger people (age 22–30 years) had an average ambient greenness of 0.26 with access to an average of 523 GBS. In contrast, older people (age 61–70 years) had a higher EVI value of 0.30, but access to an average of 382 GBS.

Differences were seen in the distribution of access to GBS and ambient exposure when grouped by CMD and rurality (*Figure 17*) with rural areas having fewer publicly accessible GBS (i.e. mainly privately owned farmland with varying rights of access).

Types of green and blue spaces

In the time-aggregated cohort, deciduous woodland was the most common nearest GBS type within 1600 m of home (*Appendix 4*, *Table 31*). This was recorded for 30% of the cohort (n = 701,966). The next nearest types were grassland, 8.5% (n = 199,458), and other grounds, 8.5% (n = 198,256). The most common nearest type of blue space was river, 5.1% (n = 129,239) followed by beach, 3.9% (n = 92,225). For individuals with a CMD, the most common nearest types of GBS were allotments and sports pitches (both 24.5% of the CMD cohort). For individuals without a CMD, the most common was deciduous woodland (n = 549,558). Living near a lido was associated with the largest proportion of the cohort without a CMD (83.3%; note, however, n = 670).

Research question 1: Is greater residential exposure and access to GBS associated with the likelihood of seeking help for a common mental health disorder in general practice?

We used the framework of our time-aggregated analysis to address each of our research questions. The results to each question are presented using pairs of tables that contain results for associations of CMD with (1) our primary exposure of ambient greenness, and (2) secondary exposure of counts of GBS access.
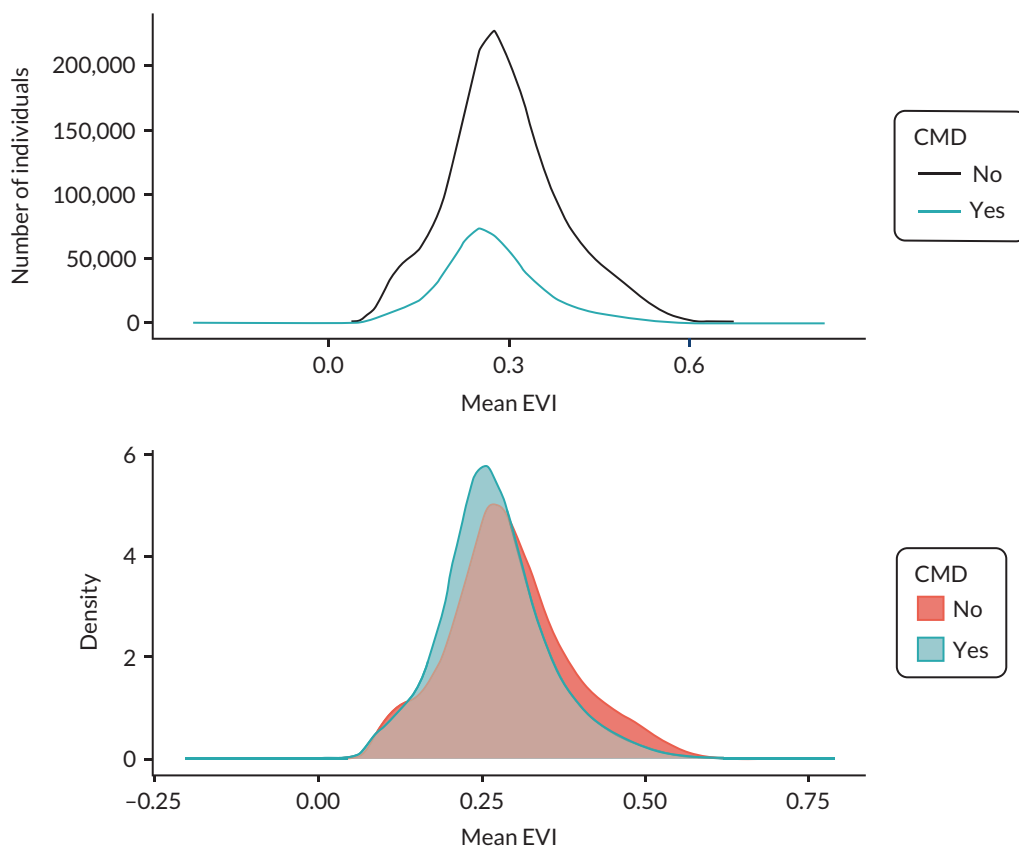
FIGURE 15 Distributions of average ambient exposure (EVI) within the time-aggregated cohort subgrouped by those with and without a CMD.
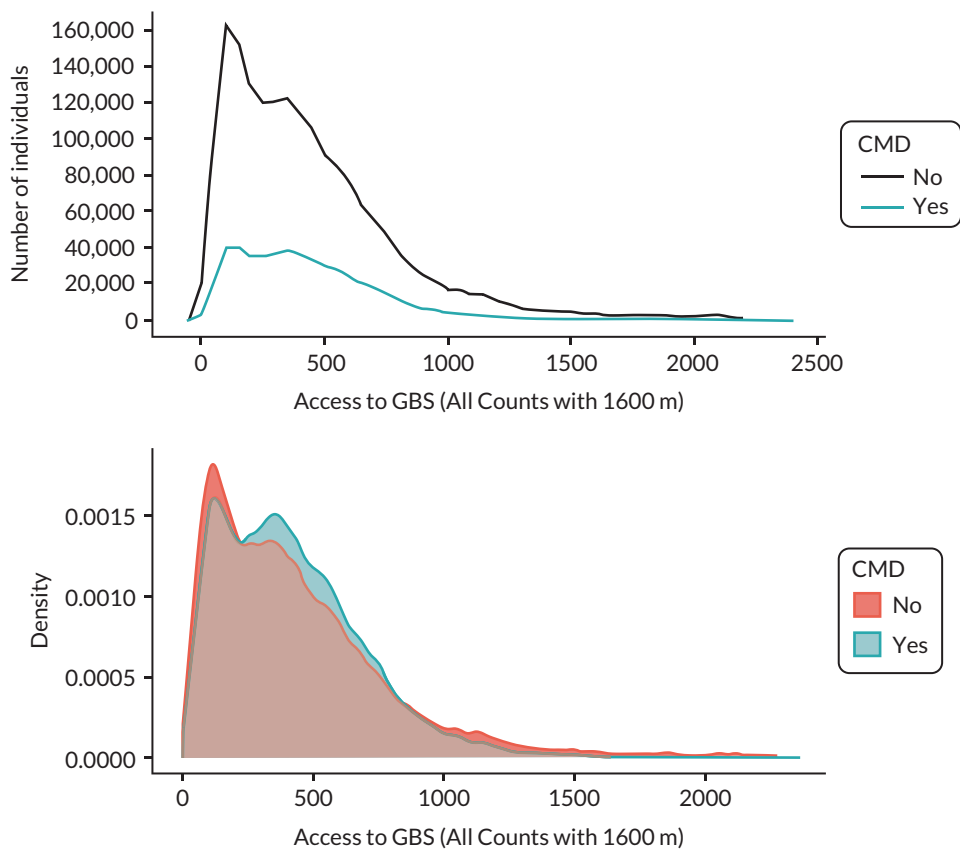
FIGURE 16 Distributions of average access to a GBS within the time-aggregated cohort subgrouped by those with and without a CMD. Density represents the density of the population as a function of exposure value within the subgrouped population (with or without a CMD).

TABLE 7 Average exposure (EVI, access to GBS) by sociodemographics for the time-aggregated cohort

Group	Ambient greenness (EVI)			Counts of GBS		
	Mean	**SD**	**IQR**	**Mean**	**SD**	**IQR**
All individuals	0.29	0.09	0.11	441.58	340.72	410.05
CMD outcome						
CMD	0.27	0.08	0.10	430.24	294.08	377.00
No CMD	0.29	0.10	0.12	444.77	352.64	419.26
Age group (years)						
16–21	0.30	0.10	0.13	491.82	392.31	436.50
22–30	0.26	0.09	0.11	523.48	408.53	458.48
31–40	0.27	0.09	0.10	453.19	332.24	407.74
41–50	0.29	0.09	0.11	412.64	298.63	387.00
51–60	0.30	0.09	0.11	401.53	296.20	389.40
61–70	0.30	0.09	0.12	382.49	285.34	377.00
71–80	0.30	0.09	0.11	395.23	290.27	385.00
80+	0.28	0.09	0.11	401.08	301.52	403.35
Sex						
Male	0.29	0.09	0.11	440.58	337.83	409.66
Female	0.29	0.09	0.11	442.62	343.69	410.54
WIMD						
1 – Most deprived	0.25	0.08	0.10	495.13	253.75	422.00
2	0.27	0.08	0.10	443.93	291.10	352.87
3	0.30	0.10	0.14	413.21	347.62	414.00
4	0.31	0.11	0.15	405.72	401.33	403.00
5 – Least deprived	0.29	0.08	0.10	448.14	382.03	422.00
Moved home						
Does not move	0.29	0.10	0.12	446.32	334.40	413.00
Moved once	0.29	0.10	0.11	432.63	360.11	413.17
Moved more than once	0.28	0.09	0.11	435.91	335.67	390.44
Rurality						
Urban	0.25	0.07	0.09	554.64	346.99	374.13
Rural	0.41	0.09	0.12	162.60	115.21	129.00
Town and fringe	0.31	0.07	0.09	219.84	130.79	180.00

WIMD is a measure of socioeconomic status; ambient greenness was measured using the EVI within 300 m of home; all GBS were counted within 1600 m of home. IQR, interquartile range; SD, standard deviation.

Primary outcome: ambient greenness

Ambient greenness around the home location was associated with reduced odds of CMD, both before and after adjustment for potential confounders. After adjustment, a 0.1 increase in standardised EVI was associated with 20% lower odds of a CMD (*Table 8*, *Figure 17*). Numerous studies have found
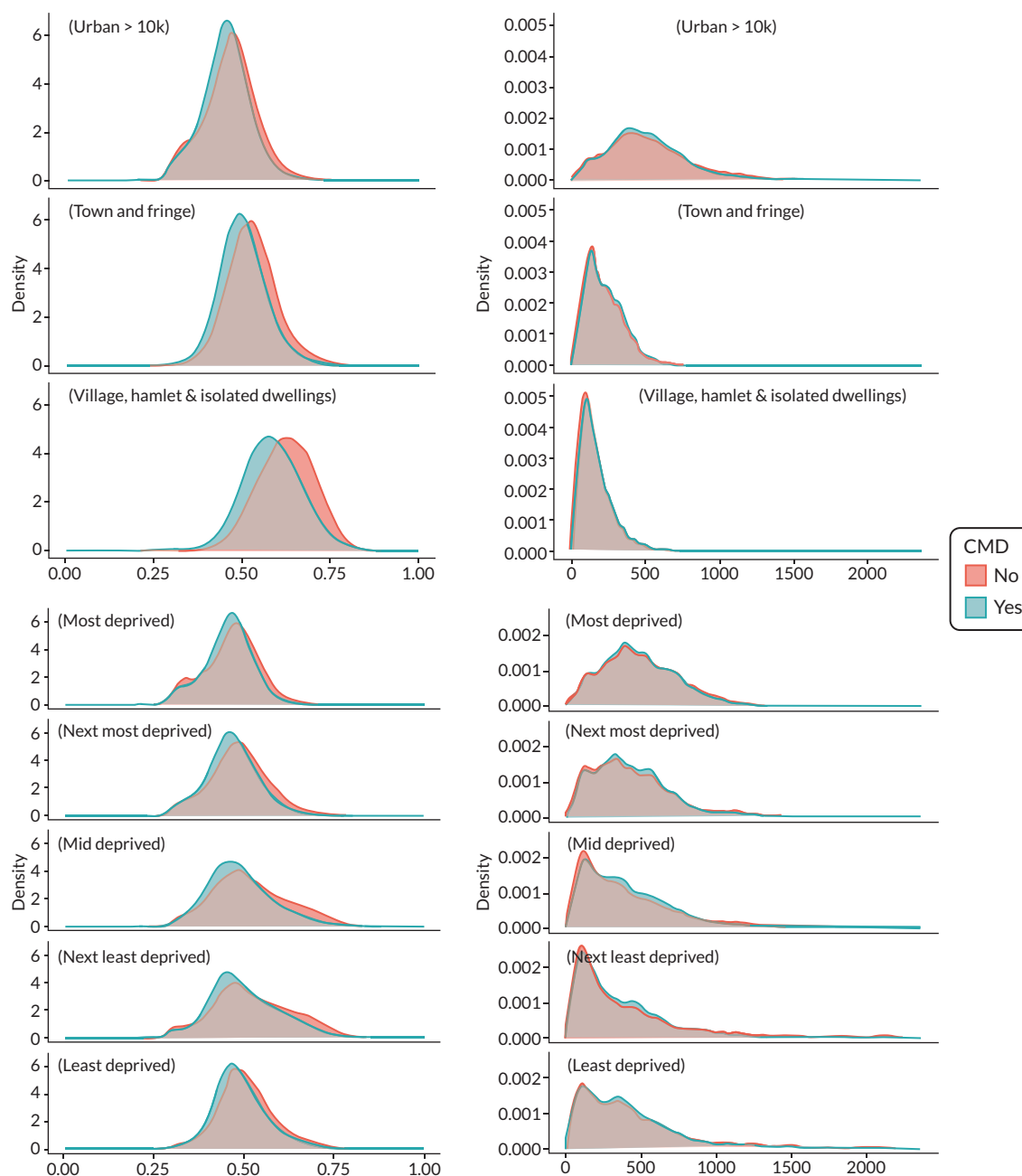
FIGURE 17 Density distributions of the subgrouped populations (with or without a CMD) population of average exposure (EVI and access to GBS) by deprivation and rurality of the time-aggregated cohort.

that satellite-derived measures of greenness are associated with better mental health. These studies have generally used NDVI as a measure of exposure to greenspace.[138,162] We also calculated NDVI[163] (*Chapter 4*) and found the same patterns as our primary exposure for ambient greenness (EVI). These results are reported in *Appendix 4*, *Table 32*.

People living in the most deprived areas had 64% higher odds of a CMD compared with those living in the least deprived areas (adjusted odds ratio, AOR, 1.64, 95% confidence interval, CI, 1.62 to 1.66; *Table 8*). The detrimental effect of living in the most versus least deprived areas is thus about three times larger than the benefit associated with a 0.1 unit increase in ambient greenness around the home. Compared with those who remained in the same home throughout their time in the study, people who moved had lower odds of a CMD (moved once AOR 0.70, moved more than once AOR 0.60, *Table 8*).

TABLE 8 Unadjusted and adjusted logistic regression model results for CMD outcome and associations with primary exposure (ambient greenness exposure; EVI)

Ambient greenness annual mean AOR for 0.1-unit increase	AOR	95% CI	
	0.80	0.80	0.81
Age group (years)			
16–21 (ref)	-	-	-
22–30	1.65	1.63	1.67
31–40	1.71	1.69	1.73
41–50	1.59	1.57	1.60
51–60	1.29	1.27	1.31
61–70	0.95	0.94	0.96
71–80	0.94	0.93	0.96
80+	0.90	0.88	0.92
Sex			
Male (ref)	-	-	-
Female	1.84	1.83	1.85
Birth in household			
No (ref)	-	-	-
Yes	0.75	0.74	0.75
Death in household			
No (ref)	-	-	-
Yes	0.65	0.64	0.65
Charlson Comorbidity Index			
	0.96	0.96	0.96
WIMD			
1 – Most deprived	1.64	1.62	1.66
2	1.38	1.37	1.40
3	1.20	1.19	1.21
4	1.08	1.06	1.09
5 – Least deprived (ref)	-	-	-
Moved home			
Did not move (ref)	-	-	-
Moved once	0.70	0.69	0.70
Moved more than once	0.60	0.59	0.60
Rurality			
Urban	0.80	0.79	0.81
Rural	1.03	1.02	1.05
Town and fringe (ref)	-	-	-

AOR based on simultaneous modelling for age, sex, household events, comorbidity, WIMD, moving home, rurality. Charlson Comorbidity Index (a score 0–33, based on 17 items) is included as a covariate. Results presented for 0.1-unit increments in EVI.

Access to green and blue spaces

The odds of a CMD decreased with greater access to GBS (*Appendix 4*, *Figure 42*). A 0.1-unit greater GBS count within 1600 m of the home (equivalent to 236 more GBS) was associated with 7% lower odds of a CMD (*Table 9*).

TABLE 9 Adjusted logistic regression model results for secondary exposures (access to GBS count) for a 0.1 unit increase in count of GBS within 1600 m

	Model: GBS counts			Model: distance to nearest GBS			Model: average GBS distance		
	AOR	95% CI		AOR	95% CI		AOR	95% CI	
Average access to GBS (all counts) AOR for 0.1-unit increase	0.93	0.93	0.93	1.05	1.04	1.06	1.02	1.01	1.03
Age group (years)									
16–21 (ref)	-	-	-	-	-	-	-	-	-
22–30	1.81	1.79	1.83	1.78	1.76	1.80	1.78	1.76	1.80
31–40	1.79	1.76	1.81	1.79	1.77	1.81	1.79	1.77	1.81
41–50	1.57	1.56	1.59	1.60	1.59	1.62	1.61	1.59	1.62
51–60	1.26	1.24	1.27	1.29	1.27	1.30	1.29	1.27	1.30
61–70	0.91	0.90	0.92	0.94	0.93	0.95	0.94	0.93	0.95
71–80	0.93	0.91	0.94	0.95	0.93	0.96	0.95	0.93	0.96
80+	0.92	0.90	0.93	0.93	0.92	0.95	0.94	0.92	0.95
Sex									
Male (ref)	-	-	-	-	-	-	-	-	-
Female	1.85	1.83	1.86	1.84	1.83	1.85	1.84	1.83	1.85
Birth in household									
No (ref)	-	-	-	-	-	-	-	-	-
Yes	0.72	0.71	0.73	0.73	0.72	0.74	0.73	0.72	0.74
Death in household									
No (ref)	-	-	-	-	-	-	-	-	-
Yes	0.63	0.62	0.63	0.63	0.62	0.64	0.63	0.63	0.64
Comorbidity index	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96
WIMD									
1 – Most deprived	1.81	1.79	1.83	1.79	1.77	1.81	1.78	1.77	1.80
2	1.44	1.43	1.46	1.45	1.43	1.46	1.44	1.43	1.46
3	1.17	1.16	1.18	1.18	1.17	1.20	1.18	1.17	1.19
4	1.03	1.02	1.04	1.05	1.04	1.06	1.05	1.04	1.06
5 – Least deprived (ref)	-	-	-	-	-	-	-	-	-
Moved home									
Does not move (ref)	-	-	-	-	-	-	-	-	-
Moves once	0.69	0.69	0.70	0.70	0.70	0.71	0.70	0.70	0.71
Moves more than once	0.60	0.59	0.60	0.61	0.60	0.61	0.61	0.60	0.61

AOR based on simultaneous modelling for age, sex, household events, comorbidity, WIMD, moving home, rurality. Charlson Comorbidity Index (a score 0–33, based on 17 items) is included as a covariate. Results presented for 0.1-unit increments in the number of potentially accessible GBS.

We explored associations with alternative GBS access metrics: distance to the nearest GBS and the average distance to GBS. Larger distance was associated with greater odds of a CMD (*Table 9*). For every 360 m (0.1 unit) increase in distance to the nearest GBS, there was 5% higher odds of a CMD (AOR 1.05, 95% CI 1.04 to 1.05). A 0.1-unit greater mean distance to GBS (625 m) was associated with 2% higher odds of a CMD (AOR 1.02, 95% CI 1.01 to 1.03).

Investigations by subgroup showed that socioeconomic status was associated with CMD (*Table 9*). Focusing on the results for counts of GBS, we found that people living in the most deprived areas had higher odds of a CMD compared with those living in the least deprived areas (AOR 1.81, 95% CI 1.79 to 1.83). There were similar patterns for both of our other access metrics. Those who moved home had lower odds of a CMD compared with those who did not move home (moved once AOR 0.69, moved more than once AOR 0.60, *Table 9*).

Predicted probabilities

Using fully adjusted model results, we predicted the probability of having a CMD as a function of each GBS exposure type, while holding all other variables constant. The predictive probabilities are based on the models detailed in *Appendix 4*, *Table 33* and *Table 34*. These are adjusted for age, sex, death and birth in household, comorbidity, deprivation, and house moves and used covariates associated with reference groups. Typical normalised EVI exposure values range from 0.25 to 0.80. Our model predicts an approximately 30% lower probability of a CMD at the bottom of this range compared with the top. Typical normalised GBS access values range from 0 to 0.55; CMD probability is approximately 10% lower at the bottom of the range compared with the top (*Figures 18* and *19*).

Research question 2: Is the association between exposure or access to green and blue spaces and seeking help for a common mental health disorder modified by socioeconomic disadvantage and moving home?

We defined subgroups within our cohort based on categories of area deprivation, history of CMD, settlement type (rurality) and moving home (migration). Through separate analyses of these subgroups,
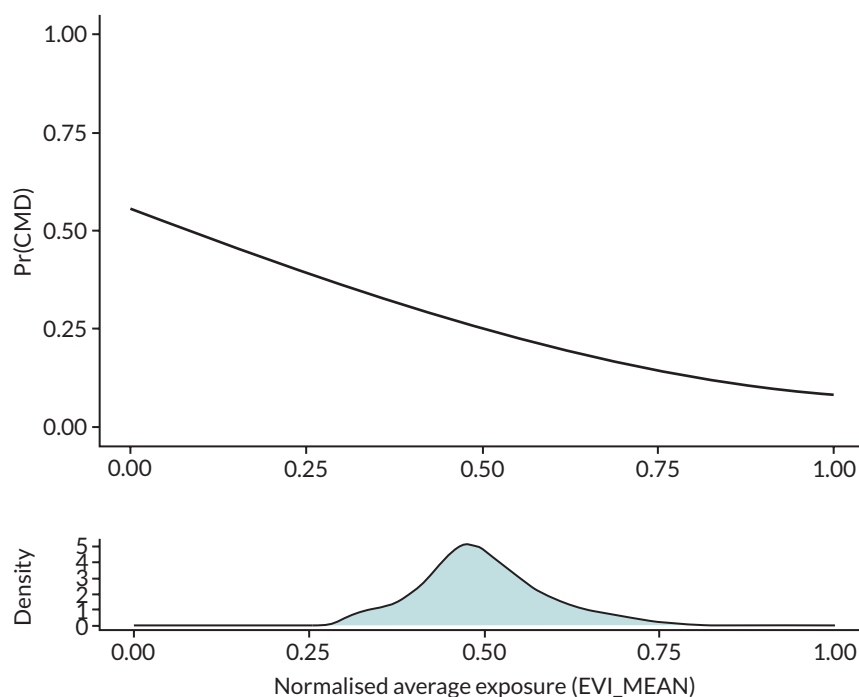
FIGURE 18 Predicted probabilities of a CMD, stratified by ambient greenness (EVI). EVI is defined as normalised annual ambient exposure to EVI within 300 m of home. GBS access is defined here as a normalised count of all GBS within 1600 m of home.
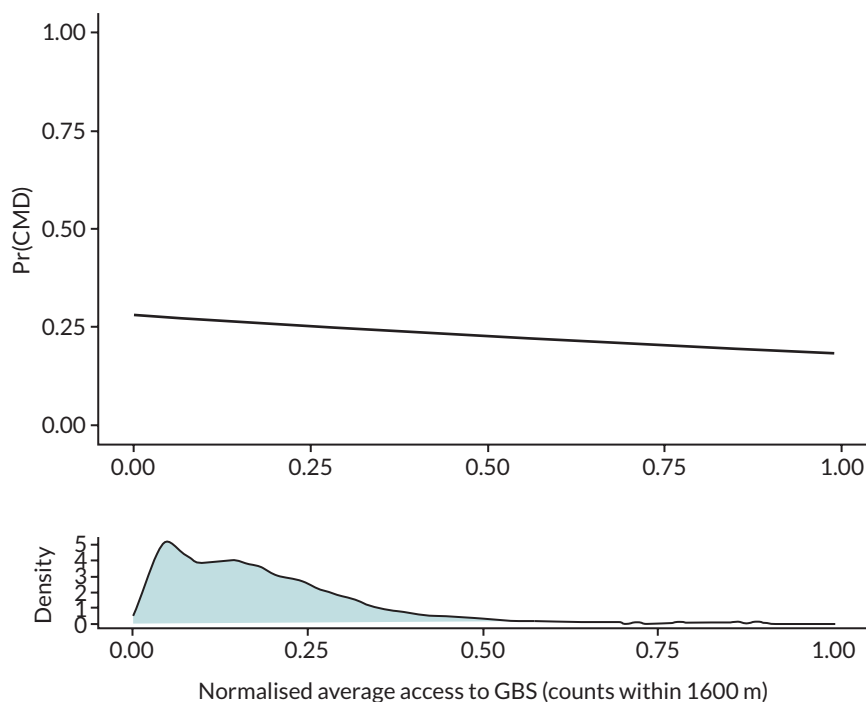
FIGURE 19 Predicted probabilities of a CMD, stratified by potential GBS access. EVI is defined as normalised annual ambient exposure to EVI within 300 m of home. GBS access is defined here as a normalised count of all GBS within 1600 m of home.

we calculated AORs for associations between ambient greenness or potential access to GBS and CMD (*Table 10*). We present AORs for 0.1-unit increments in EVI and in GBS count.

Deprivation

When stratified by deprivation quintile, both ambient greenness and access to GBS were associated with CMD across all quintiles (*Table 10*). However, we found no modifying effect of deprivation for increasing ambient greenness; there was no clear trend across deprivation categories of the magnitude of association between EVI and CMD. In contrast, we found the association between access to GBS and CMD to be modified by deprivation: those from more deprived groups benefited most from greater access to local GBSs (AOR for most deprived areas 0.90, 95% CI 0.90 to 0.91; AOR for those living in the least deprived areas 0.94, 95% CI 0.94 to 0.95; *Table 10*).

Settlement type

Associations between GBS exposure and access are modified by rurality. Increasing ambient greenspace reduced CMD likelihood for all groups, but the greatest reduced risk was in rural communities (AOR 0.59, 95% CI 0.58 to 0.59). The relationship between GBS access and CMD varied by settlement type. In urban areas, greater access was associated with lower odds of having a CMD. The opposite was seen for those living in non-urban areas, n greater GBS access (town and fringe; AOR 1.05, 95% CI to 1.03,1.06; village, hamlet and isolated dwellings: rural; AOR 1.10, 95% CI 1.08 to 1.12) was associated with higher odds of having a CMD (*Table 10*).

Moved home

We found a small modifying effect of moving home on the relationship between counts of GBS and CMD. Greater GBS access was associated with lower CMD likelihood among those who moved home at least once (moved once: AOR 0.91, 95% CI 0.91 to 0.92; moved more than once: AOR 0.89, 95% CI 0.89 to 0.90; did not move: AOR 0.94, 95% CI 0.93 to 0.94). Ambient greenness was associated with lower odds of a CMD in all 'moving home' groups but the effect was strongest for those who did not

TABLE 10 Adjusted logistic regression model results stratified by deprivation, age, birth in house, death settlement type, moving home, and historical CMD diagnosis for a 0.1-unit increase in primary (EVI) and secondary exposures (access)

	(%)	(AOR)	Ambient greenness, EVI			Counts of GBS		
			(AOR)	(95% CI)		(AOR)	(95% CI)	
WIMD[a]								
1 – Most deprived	20.2		0.86	0.85	0.87	0.90	0.90	0.91
2	19.8		0.78	0.78	0.79	0.92	0.92	0.93
3	20.5		0.77	0.77	0.78	0.94	0.94	0.95
4	19.0		0.83	0.82	0.83	0.93	0.92	0.93
5 – Least deprived	20.5		0.77	0.77	0.78	0.94	0.94	0.95
Settlement type[b]								
Urban	68.8	1.01	0.85	0.84	0.85	0.89	0.89	0.89
Town and fringe	15.2	1.0	0.59	0.59	0.60	1.05	1.03	1.06
Rural	16.0	0.73	0.59	0.58	0.59	1.10	1.08	1.12
Moved home[c]								
Does not move	61.5	1.0	0.77	0.77	0.78	0.94	0.93	0.94
Moves once	16.3	0.75	0.86	0.86	0.87	0.91	0.91	0.92
Moves more than once	22.2	0.71	0.92	0.91	0.93	0.89	0.89	0.90
CMD history[d]								
Diagnosis before entering study	10.1	6.78	0.68	0.68	0.68	1.00	0.99	1.00
No diagnosis before entering study	89.9	1.0	0.84	0.84	0.85	0.91	0.91	0.92

a Models stratified by WIMD were adjusted for age, sex, birth/death in household, comorbidity index and moving home. Count of GBS AOR for a 0.1-unit increase in count of GBS within 1600 m.
b Models stratified by settlement type were adjusted for age, sex, birth/death in household, comorbidity index, WIMD and moving home. Count of GBS AOR for a 0.1-unit increase in count of GBS within 1600 m.
c Models stratified by moving home were adjusted for age, sex, birth/death in household, comorbidity index and WIMD. Ambient greenness AOR for a 0.1-unit increase in EVI. Count of GBS AOR for an 0.1 unit increase in count of GBS within 1600 m.
d Models stratified by CMD history were adjusted for age, sex, birth/death in household, comorbidity index and WIMD. Ambient greenness AOR for a 0.1-unit increase in EVI. Count of GBS AOR for an 0.1 unit increase in count of GBS within 1600 m.

move home (AOR 0.77, 95% CI 0.77 to 0.78) compared with those who moved (moved once: AOR 0.86, 95% CI 0.86 to 0.87; moved more than once: AOR 0.92, 95% CI 0.91 to 0.93).

Common mental health disorder history

Although we stopped follow-up for this time-aggregated cohort when the first CMD was recorded within the study period, we were interested in differences between people who had an historical CMD before entering the cohort and those who did not. We therefore stratified the analysis by individuals who had an historical CMD before 2008. There was a larger association with ambient greenness for people with an historical CMD (AOR 0.68, 95% CI 0.68 to 0.68) in comparison with those without (AOR 0.84, 95% CI 0.84 to 0.85).

The effect of a 0.1-unit greater GBS access on lowering the odds of having a CMD, was significant for those without an historical CMD (AOR 0.91, 95% CI 0.91 to 0.92). For those with a historical CMD, there was no association (AOR 1.00, 95% CI 0.99 to 1.00).

Variable	OR	P Value	CI	Count (%)
Nearest type of GBS				
Park (ref)				21920 (0.94)
Allotments	1.19	< 0.001	(1.14, 1.24)	47933 (2.05)
Lido	0.78	0.02	(0.64, 0.97)	670 (0.03)
Mixed green space	0.93	0.12	(0,85, 1.02)	4241 (0.18)
Harbour/dock/marina	0.95	0.07	(0.9, 1)	14411 (0.62)
Religious grounds	0.95	< 0.001	(0.92, 0.98)	143270 (6.12)
Canal	0.97	0.65	(0.85, 1.11)	1601 (0.07)
Multiple blue spaces	0.97	0.09	(0.93, 1)	111426 (4.76)
Estuary	0.98	0.87	(0.75, 1.28)	365 (0.02)
Other grounds	0.98	0.23	(0.95, 1.01)	198256 (8.47)
Pond	1	0.82	(0.97, 1.04)	84978 (3.63)
Botanical gardens	1.02	0.8	(0.89, 1.16)	1581 (0.07)
Coniferos	1.03	0.09	(1, 1.07)	77667 (3.32)
Lake	1.03	0.61	(0.93, 1.13)	2928 (0.13)
Marsh	1.03	0.24	(0.98, 1.09)	13623 (0.58)
Recreational	1.04	0.12	(0.99, 1.1)	16062 (0.69)
Beach	1.05	0.02	(1.01, 1.09)	92224 (3.94)
River	1.05	0.01	(1.01, 1.08)	129236 (5.52)
Grassland	1.06	< 0.001	(1.03, 1.1)	199479 (8.52)
Undefined	1.06	0.62	(0.83, 1.36)	459 (0.02)
Moor heath	1.07	< 0.001	(1.03, 1.1)	163899 (7)
Deciduous	1.08	< 0.001	(1.05, 1.12)	701959 (29.98)
School grounds	1.08	< 0.001	(1.04, 1.12)	160661 (6.86)
Reservoir	1.09	0.04	(1, 1.18)	4209 (0.18)
Play areas	1.12	< 0.001	(1.08, 1.16)	105347 (4.5)
Quarry	1.16	< 0.001	(1.08, 1.25)	5109 (0.22)
Sports pitches	1.17	< 0.001	(1.13, 1.22)	37547 (1.6)
Meadow	1.31	0.01	(1.07, 1,59)	530 (0.02)

FIGURE 20 Adjusted logistic regression model results for nearest type of GBS compared with park as nearest GBS. Models adjusted for age, sex, birth/death in household, WIMD, comorbidity index and moving home.

Types of green and blue spaces

Individuals with specific types of nearest GBS including play areas, quarries, sports pitches, and allotments were associated with 12% to 19% greater odds of CMD compared with those living nearest to parks (*Figure 20*). In contrast, people living nearest to religious grounds were associated with 5% lower odds of CMD relative to people living closer to parks.

Discussion

Main findings

Question 1: Do people with different GBS exposure through time have different associated risks of having a CMD?

We found that ambient greenness (EVI) and all measures of GBS access were associated with lower OR of having a CMD. The strongest association was seen for ambient greenness compared to more potentially accessible GBS (20% lower odds of seeking help for CMD compared with 7%, respectively).

Question 2: Is the association between changes in exposure or access to GBS, and CMD modified by multiple sociophysical variables, migration, and socioeconomic disadvantage?

When we stratified our models by area-level deprivation indices, we found that people were less likely to have a CMD when they lived in a home surrounded by more ambient greenness and had increased access to GBS, across deprivation quintiles. There was some variability, but no pattern in EVI–CMD associations across deprivation quintiles. However, there was some suggestion that the inverse association between GBS counts and CMD was stronger in more deprived areas. Focusing on the results for counts of GBS, we found that people living in the most deprived areas had higher odds of a CMD compared with those living in the least deprived areas (AOR 1.81, 95% CI 1.79 to 1.83). The ambient greenness–CMD association was stronger among people who had not moved home during their time in the cohort than among those who had not moved. Ambient greenness was associated with lower odds of a CMD in all 'moving home' groups but the effect was strongest for those who did not move home (AOR 0.77, 95% CI 0.77 to 0.78) compared with those who moved (moved once: AOR 0.86, 95% CI 0.86 to 0.87; moved more than once: AOR 0.92, 95% CI 0.91 to 0.93). The magnitude of association between ambient greenness and CMD increased around two-fold in people with an historical CMD (32% lower odds for a standard deviation increase in EVI vs. 16% for those without).

When stratified by rurality, we found that increasing ambient greenness is beneficial to people living in both urban and rural areas. However, increased access to GBS was only significant for individuals living in urban areas. In town and fringe and rural areas, a greater number of GBS within 1600 m was associated with higher odds of having a CMD. For individuals living in urban areas, there were reduced odds of having a CMD with greater EVI.

Comparing people whose nearest GBS was a park with those nearest to another type of GBS gave mixed results. Some comparisons indicated a higher likelihood of having a CMD, while others indicated a lower likelihood. It should be noted that these do not indicate the absolute association between these different types of GBS and CMD. However, this analysis suggests that different qualities of GBS have different impacts on CMD outcomes and that not all GBS are equally beneficial.

Wider context
One fifth of our time-aggregated cohort had a CMD, as did more females than males. This is consistent with survey results published on the Welsh population.[164] For adults aged 16 years and over living in Wales, we found that living in an area with greater ambient greenspace was inversely associated with having a CMD. This is consistent with previous studies in the Netherlands and Spain that found greenness in the immediate environment around the home location supported good mental health.[160,165]

Gascon et al.[166] recommended that satellite-derived measures of greenness should be interpreted with care. Our ambient greenness exposure is based on satellite-derived EVI values, where raw values from 0.2 to 0.8 represent healthy vegetation. A larger EVI score does not necessarily equate to more greenness by area but instead represents a larger volume of green; there is increased biomass. A small forest, for example, could produce the same EVI score as a large grass-covered area. Combining multidimensional aspects into a single EVI figure makes the relative qualities of greenspace difficult to interpret. This echoes the main conclusion of a 2018 systematic review.[82] Future research should consider the function of a space and dimensions such as biodiversity.

Greater potential access to GBS was also associated with better metal health for all metrics of GBS. Again, this finding is consistent with previous studies where poorer access to GBS (defined by distance or travel time) was associated with worse mental health outcomes.[138,167,168] It is useful to have this finding confirmed based on household measures of GBS access. Our study is based on potential to access GBS but studies have found that people tend to engage with GBS within 1600 m of their home.[169] Some GBS may be appropriate to most of the local population, however, the type of space may be most important to some subgroups. Evidencing the benefits of GBS types is likely to be helpful for policy makers and

planners designing policies and interventions to include vulnerable groups who would benefit the most from these spaces. Current WHO indicators are based on simply having a half-hectare GBS accessible within 300 m of home, without thought of the type of space needed.[114,170] This study sheds new light on which types of GBS may be more important, with the likelihood of a CMD varying by type of nearest GBS. This builds on previous research that states that the relationship between GBS and mental health is complex, with the specific function of a GBS important to individuals.[171]

Our results for different population subgroups suggest that future research should investigate the reasons why each subgroup is impacted differentially by greenness exposure and potential GBS access. This agrees with the work of others that there is not a 'one size fits all' green space, with different population groups likely to benefit from specific GBS tailored to their needs. Our work provides evidence toward the development of more specific public health policies and interventions.

Modification by multiple sociophysical variables, moving home, and socioeconomic disadvantage

Our stratified analyses produced some novel results. Prior to this study, there was limited evidence on how relationships vary sociodemographically, although a number of studies do suggest that people in lower socioeconomic groups and those living in more socioeconomically deprived areas may benefit disproportionately from access to GBS (the 'equigenesis' hypothesis).[72,172] We find no strong evidence that the association between exposure or access to GBS and CMD is modified by deprivation. However, there was some suggestion that the preventive association we observed between GBS counts and CMD is stronger for people living in areas of higher deprivation. Our results also suggest that the association between exposure or access to GBS and CMD were modified by a prior CMD, rurality and moving home.

Deprivation

For people living in the least deprived areas, there were 6% lower odds of seeking GP care for CMD if their potential to access GBS was higher. Similarly, the most deprived had 10% lower odds of CMD-related help seeking, indicating the greater potential benefits to mental health for people living in deprived areas but with greater accessible GBS. Greater ambient greenness reduced CMD likelihood across all deprivation strata. However, this time the pattern was reversed, with the lowest odds of seeking care for CMD for people living in the least deprived group. This reversal may indicate that greater access to (smaller) parcels of diverse land use, as used in the GBS access metric, are more conducive to maintaining mental health, compared to overall ambient greenness. Without a measure of quality, we do not know if areas of ambient greenness include larger areas of inaccessible grass or derelict scrubland. Deprivation has been previously documented as modifying the relationship between greenspace and mental health.[173,174] It has been proposed that the benefits of a greenspace are not equally experienced by individuals, but that the characteristics of the population using the greenspace are important to consider.[175] The use of material deprivation at the household level is presented in *Chapter 7* for those in the survey subsample to contrast with area-level deprivation.

Historical common mental health disorders

Our results suggest that the relationships between exposure and access to GBS and CMD likelihood are modified by an individual's CMD history. For individuals with an historical CMD before entering the cohort, larger ambient exposure was more beneficial than for those who without a history of CMD. This is a potentially important point for reducing mental health inequality. For individuals with an historical CMD, those with 0.1-unit higher ambient exposure to greenness were 32% less likely to have a CMD. In contrast, those without an historical CMD were 16% less likely to have a CMD for the same increment in exposure. This suggests that exposure to ambient greenness has different impacts on people for prevention compared to those who experience restorative benefits.

Increasing access to GBS benefited those without an historical CMD but no effect was found for those with a historical CMD. We purposefully set up this time-aggregated investigation to focus on preventative effects, thus we cannot show whether increasing ambient greenness works as a treatment for CMD. However, we evidence that improving exposure and access to GBS is beneficial to those who may be more susceptible to CMDs, through the reduced likelihood that they will seek help from a GP.

Rurality

Previous studies have generally focused on urban GBS access,[176–180] but our study stratified individuals living in all types of settlements, from rural, to towns, and urban morphologies. When stratified, we found that more ambient greenness may be beneficial to people living in both urban and rural areas. People living in urban areas had 11% reduced odds of seeking help for a CMD with greater GBS access. In rural areas, people living in similar GBS accessible locations, had 10% increased odds of a CMD. Investment in urban greenspaces is important but rural communities should not be forgotten. It may be that guided nature-based activities are needed to transform the potentially accessible GBS into a benefit for mental health in rural areas. Compared with urban communities, individuals in rural communities can be more susceptible to loneliness,[181] have less social connectedness,[182] and have poorer access to healthcare.[183,184] Further research is needed to better understand what intervention activities and mechanisms in GBS promote good mental health and well-being for people living in rural communities.

Moving home

We were interested in those who moved home, because these people were more likely to experience large changes in exposure and access to GBS. We found a potentially modifying effect for both ambient greenspace and access to GBS, but in opposite directions. Higher ambient greenness was associated with a greater reduced odds of a CMD in those who did not move. Conversely, individuals who moved more than once were less likely to have a CMD with higher GBS. In a longitudinal study, Alcock *et al*.[45] found that individuals who moved to greener areas had sustained better mental health but no previous research has accounted for moving home in a population-level study. Our next chapter will discuss the longitudinal impacts of moving home.

Strengths and limitations

This analysis has several methodological strengths including linking national environmental data and with population-wide anonymised health records. Investigating such a large population enables us to better understand population-level trends that can contribute to evidence-based policy and intervention development (see *Chapters 2* and *3*). Our results distinguish between the immediate home environment and the potential to access GBS in the nearby neighbourhood. This methodological choice acknowledges that there are different ways that people will interact and engage with GBS in their day to day lives. Furthermore, studies typically analyse data on a small-area basis. Our analysis has linked data at a higher spatial resolution which will minimise the modifiable areal unit problem,[185] reduce spatial smoothing and the degree of ecological fallacy that occurs when data are aggregated over larger statistical areas. Careful data preparation ensured that the CMD outcome always followed exposure, reducing the liability for circularity in this time-aggregated design.[186,187] We included all adults in Wales, only censoring exposure and follow up for those who had a recorded CMD, thereby reducing selection bias. Our individual level approach reduced the potential that we detected health improvements only due to the replacement of the population due to gentrification.[188] Our inclusion of certain confounders and mediators that have been suggested in previous literature sheds new light on sociodemographic influences on the relationship between GBS and CMDs.

The study has limitations. Firstly, identifying 'real' CMD cases requires professional clinical consideration of patients' medical history and personal circumstances, and these may be imperfectly captured in primary healthcare records.[189] We introduced the algorithm used to detect CMD in *Chapter 2*, capturing

the strengths and limitations of this method in detail. Routinely collected data, survey data and clinician diagnoses have a different set of biases. For example, with self-reported data, people may report closer to the perceived social norms, thereby underestimating the prevalence of a condition.[190] This algorithm captures recorded symptoms, as reported to a GP, and associated treatments and/or diagnoses made for the patient. Secondly, those who have recovered from a CMD episode or experience a very severe case might not appear in a general practice record. Therefore, we acknowledge that at either end of the CMD symptom severity scale the algorithm may have lower specificity.

The estimation of exposures to greenspace using vegetation indices is widely described in the literature.[191–194] However, there is no doubt that an EVI score (and all vegetation scores) is influenced by a range of atmospheric, vegetation type and seasonal parameters which may have limited our ability to estimate EVI exposures at peak greenness each year. Given this, we are inclined to treat EVI scores as indicative scores of greenness around a household rather than absolute.

Despite these limitations, the exploratory time-aggregated analysis described here provides a good basis for the longitudinal analyses described in the following chapter.

Chapter 6 Green and blue spaces around the home and common mental health disorders: a retrospective cohort with assessment of exposure change

Introduction

Growing evidence suggests that living close to GBS (e.g. parks, gardens, trails, ponds, lakes, rivers) is associated with positive health outcomes including physical and mental health and child development.[87] However, there is relatively limited robust, longitudinal evidence on whether changes in GBS affect future risk for CMD.[8,45,82] This chapter explores whether changes in ambient greenness around the home and/or access to GBS are associated with changes in seeking help for CMD in general practice, using the total adult population longitudinal cohort described in *Chapter 2* (*n* = 2,801,482).

Aims

We examined the research questions:

- Is a greater residential exposure and access to GBS associated with the likelihood of seeking help for a CMD in general practice?
- Is the association between exposure, or access to, GBS and seeking help for a CMD modified by socioeconomic disadvantage and moving home?
- Is a greater residential exposure associated with the likelihood of using GP services?

Methods

We completed a series of analyses for each of our outcomes (seeking help for a CMD in general practice, GPE), and our primary exposure (ambient greenness within 300 m of the home) and secondary exposure (the number of potentially accessible GBS within 1600 m of the home). People could change classification between 'seeking help for a CMD in general practice' and 'not' in each quarter of the longitudinal dataset. Equally, they could accumulate different counts of GPE in each quarter.

Outcomes

Primary outcome: common mental health disorders
The primary outcome was likelihood of seeking GP help for CMD , determined using the algorithm described in *Chapter 2* to produce a binary outcome. The algorithm identified people seeking GP help for a CMD each quarter using read codes in their health records held in the WLGP dataset (*Report Supplementary Material 1*).[103] People were defined as 'seeking help' in a particular quarter when historical diagnosis(es) were treated, and/or the GP recorded current diagnoses or symptoms (treated or untreated). A CMD drug treatment was identified as at least one prescription for an antidepressant, anxiolytic or hypnotic in the previous 12 months.[103] This same algorithm is used throughout this study,

however, in these longitudinal analyses, individuals could have periods of seeking, and of not seeking help for a CMD in general practice.

Secondary outcome: general practice events

The total number of GPE each day were available for each person in the WLGP dataset. 'Events' can include general practice interactions (e.g. with a practice nurse, or a GP) as well as investigations/tests, prescribed medication and referrals. These events reflect the general need for primary health services. We converted the daily count of GPE to a binary variable (yes/no on any given day) and then aggregated to quarterly counts (minimum = 0 if no events in the quarter, maximum ~120 if events occurred every day in the quarter). Daily aggregation of events prevented multiple test results on one day being given more 'weight' compared with interactions with a nurse or GP.

Environmental exposures

Changes in the environmental metrics could occur at the environment level (e.g. a change in the GBS itself such as building on a local park, reduction, or a new community garden on vacant ground, gain) or at the individual level, such as through moving home.[83] We obtained start and end dates of periods lived at each address for all cohort participants to enable individual level analysis. The dates allowed us to allocate dynamic measures of the environmental exposures specific to all individuals living in the household, for the relevant quarters.

Primary environmental exposure: ambient greenness

The primary environmental exposure for these longitudinal analyses was change in ambient greenness within 300 m of each residence, as measured by change in annual mean EVI (*Table 11*) compared with the previous year and baseline. Changes in annual mean EVI were also compared with the previous quarter only for people who moved home, updating mean EVI in the quarter that they moved. Change in mean EVI was estimated by subtracting an earlier quarterly mean EVI value from a later quarterly value. More complex smoothing approaches, as used in previous studies of change in alcohol outlet density, were not possible as the EVI data were not temporally rich enough.[108] EVI can range from −1 to 1 and values between 0.2 and 0.8 represent healthy vegetation. We found that mean EVIs within 300 m of a home in Wales between 2008 and 2019 fell in the range 0.10–0.62 (*Chapter 5*, *Figure 15*).

Secondary environmental exposure: potential access to green and blue spaces

The secondary environmental exposure was the potential for an individual to access GBS along a network of paths and roads, measured as the total count of GBS within 1600 m of each home (*Table 11*). We refer to this as change in 'the number of potentially accessible GBS'. While data were available for 2012 and 2018, the database behind the OSMM Topography Layer is part of a continuous improvement programme. This resulted in land use classification changes that did not reflect losses or gains in GBS (see *Chapter 3*). Thus, these data do not reliably reflect *change* in GBS over time. As a result, investigation of the association between change in the number of potentially accessible GBS and our outcomes in these longitudinal analyses were restricted to those who moved home where we allocated 2018 access measures. We grouped those who moved home once, and those who moved more than once separately.

Change in environmental exposure

Change in GBS (for both EVI and GBS access) was modelled as a continuous variable, with change relative to previous quarter, previous year and baseline (defined as the first quarter an individual enters the cohort). For EVI, we present results for increments of 0.1-unit change. A 1-unit change, typically used to present results from regression analyses, is an extremely large change in EVI, effectively representing a change from no ambient greenness to almost total coverage with highly reflective healthy

TABLE 11 Definition of environmental exposures

Environmental metric	Definition	Measure (source)	Description	Change relative to previous quarter, previous year, and baseline
Primary exposure: change in ambient greenness around the home	The presence of GBS (e.g. parks, domestic gardens) and green infrastructure (e.g. street trees, green roofs) within a 300 m linear (crow flies) buffer of each home, without assumptions on availability for public use	Annual mean EVI for each household. (Landsat satellite imagery 2008–2019)	Measure of vegetation within 300 m of the home. Measurements derived from remotely sensed satellite images. See *Chapter 3* for more detail	Continuous: 0.1-unit increase in regression analyses Categorical: large positive change: ≥ +0.15 mean EVI, no/small change (< 0.15 EVI increase or decrease), large negative change: ≥ −0.15 EVI
Secondary exposure: change in potential access to GBS around the home	The number of GBS that people can potentially access a footpath or road-accessible network buffer within 1600 m of their home	Count of GBS within 1600 m of each household location using a network model. Various vector data; see *Chapter 3* for more detail	Measure of density of potentially accessible GBS within 1600 m of the home. See *Chapter 3* for more detail	Continuous (1-unit increase in regression analyses)

vegetation. This level of change in EVI is not achievable in most local environments. For those who did not move home, we modelled change in EVI as a categorical variable (large positive change: ≥ +0.15 change in mean EVI, no/small change (< 0.15 change (increase or decrease) in mean EVI), large negative change: ≥ −0.15 change in mean EVI) to examine changes in EVI around the home that could not be attributed to moving.

Covariates

Covariates included sex (male, female), age group (years, continuous), deprivation (quintiles of Welsh IMD), home moves (yes, no), birth(s) in household (yes, no), death(s) in household (yes, no), Charlson Comorbidity Index (continuous).[195] In preliminary analyses, we also included rurality ('village, hamlet and isolated dwellings', 'town and fringe', 'urban > 10 k', defined using ONS settlement type categories).[96] However, this could represent another measure of the primary environmental exposure, EVI, so rurality was excluded from our final models to avoid overadjustment.[196] Covariates used in each analysis are listed in the relevant table.

Statistical analyses

This longitudinal analysis was carried out using a panel design with repeated observations for each participant. A multilevel model structure with quarterly observations over time (level 1) nested within individuals (level 2) allowed us to take account of clustering of observations over time using random effects. This structure also enabled the handling of unbalanced data, where the number of observations can vary for individuals, which is an artefact of a dynamic cohort.[104] It was not possible to nest within LSOAs, as individuals could change LSOA when moving home, resulting in LSOA multi-membership. For these longitudinal analyses, we tested the association of change in ambient greenness and in potential GBS access around the home through time with CMD and GPE. We did not incorporate spatial differences between people at a single point in time into our longitudinal analyses.

For the outcome CMD, we used multivariate, multilevel logistic regression models. Where GPE days was the outcome, we used multivariate, multilevel Poisson regression models to reflect the

count nature of the outcome data. For ease of interpretation, Poisson regression coefficients were exponentiated to incidence rate ratios (IRRs). We first explored associations between potential confounders and our exposures and outcomes separately. Covariates associated with both the outcome and exposures were added in a stepwise fashion to the multivariate models. The step-by-step iterative construction of the multivariate models involved adding potential explanatory variables in succession and testing for statistical significance after each iteration. Stratified analyses were conducted by deprivation and moving home. For deprivation, stratified analyses considered those living in urban areas, with the multivariate, multilevel regression models run separately for two WIMD quintiles of area-level deprivation, the least and the most deprived. For our primary outcome (EVI), stratified analyses by moving home were conducted using three categories of the number of home moves (none, one, more than one). For our secondary outcome (potential to access GBS), where individuals could only experience a change in access (in our data) by moving home (section 0), stratified analyses by moving home were conducted using two categories of the number of home moves (one, more than one). We calculated the variance partitioning coefficient for a null model without covariates for those who did not move LSOA, confirming negligible variance in our primary outcome at the LSOA level (< 1%).

We also conducted preliminary modelling of the potential impact of changes in ambient greenness on NHS costs associated with GPE via the impact on the Welsh GP budget over our study period (2008–2019). We used the total spend on GPs to estimate the Welsh general practice budget. This includes the reimbursement of drugs dispensed in general practices in Wales in 2017/18, totalling £543.9 million.[197]

Results

Cohort description

Everyone aged 16 years and over with at least one address recorded and registered with a general practice providing records to the SAIL databank between January 2008 and October 2019 (the 'study period') were included. The cohort for these longitudinal analyses was 2,801,483 individuals (female: 51%) with a total of 99,682,902 observations (see *Chapter 2* for cohort inclusion flow chart).

As described in *Chapter 2*, the cohort was broadly representative of the Welsh adult population compared with 2011 Census figures.[113] We identified 816,242 individuals as seeking help for CMD in general practice at least once during the study period (29.1% of the cohort: of those, 62.1% were female). One in five people aged 16–21 years sought help for CMD in general practice at least once between 2008 and 2019 (*Table 12*; see also *Appendix 5*, *Table 35*), the highest percentage of any age group. The percentage of those who had sought help for CMD in general practice at least once was lower in each successive age group. There was a social gradient in seeking help for CMD, with 17.2% in the least and 25.8% of residents in the most deprived areas having sought care for CMD in general practice. The mean number of GPE days per person per quarter was 3.8 (median 2, SD 4.6) quarter increased by 1 day, from 3.3 to 4.3 days.

Description of change in environmental metrics

Change in enhanced vegetation index

The primary environmental exposure was change in mean ambient greenness within a 300 m buffer, centred on each residential location. For brevity we refer to EVI, rather than ambient greenness here. Overall, there is very little or no change in EVI compared with the previous quarter for most of the cohort (mean change 0.00; *Figure 21*, left-hand side). Change in EVI compared with the previous year

TABLE 12 CMD and GPE days by cohort characteristics (at baseline) and by year

Individuals (observations)	CMD at least once between 2008 and 2019		GPE days per quarter (mean n)		
2,801,483 (91,746,917)					
	n	%	Mean	Median	SD
All	816,242	29.1	3.8	2	4.6
Sex					
Male	309,295	37.9	2.7	1.0	3.7
Female	506,947	62.1	3.5	2.0	4.0
Age group (years)					
16–21	165,784	20.3	1.6	1.0	2.3
22–30	130,716	16.0	2.2	1.0	2.8
31–40	139,257	17.1	2.4	1.0	3.2
41–50	138,900	17.0	2.8	2.0	3.5
51–60	103,420	12.7	3.8	3.0	4.0
61–70	73,960	9.1	5.1	4.0	4.5
71–80	43,048	5.3	6.4	6.0	5.1
81 +	21,157	2.6	7.2	6.0	5.6
WIMD					
1 – Most deprived	205,810	25.8	3.2	2.0	4.0
2	170,867	21.4	3.2	2.0	4.0
3	155,165	19.4	3.2	2.0	3.9
4	130,218	16.3	3.1	2.0	3.8
5 – Least deprived (ref)	137,015	17.2	2.9	2.0	3.7
New case of CMD					
Year					
2008	-	-	3.3	2	4.2
2009	65,391	2.3	3.5	2	4.3
2010	58,569	2.1	3.5	2	4.4
2011	54,384	1.9	3.6	2	4.5
2012	51,676	1.8	3.7	2	4.6
2013	50,841	1.8	3.8	2	4.6
2014	50,519	1.8	3.9	3	4.6
2015	49,103	1.8	3.9	3	4.7
2016	48,434	1.7	4.1	3	4.8
2017	47,060	1.7	4.2	3	4.9
2018	47,841	1.7	4.2	3	4.9
Jan–Oct 2019[a]	24,483	0.9	4.3	3	4.9

2008 data on new cases of CMD not presented as in the first year of the cohort these data would include treatment of historical diagnoses.
a New cases of CMD in 2019 only captures 9 months of data available to the cohort, up to October 2019.
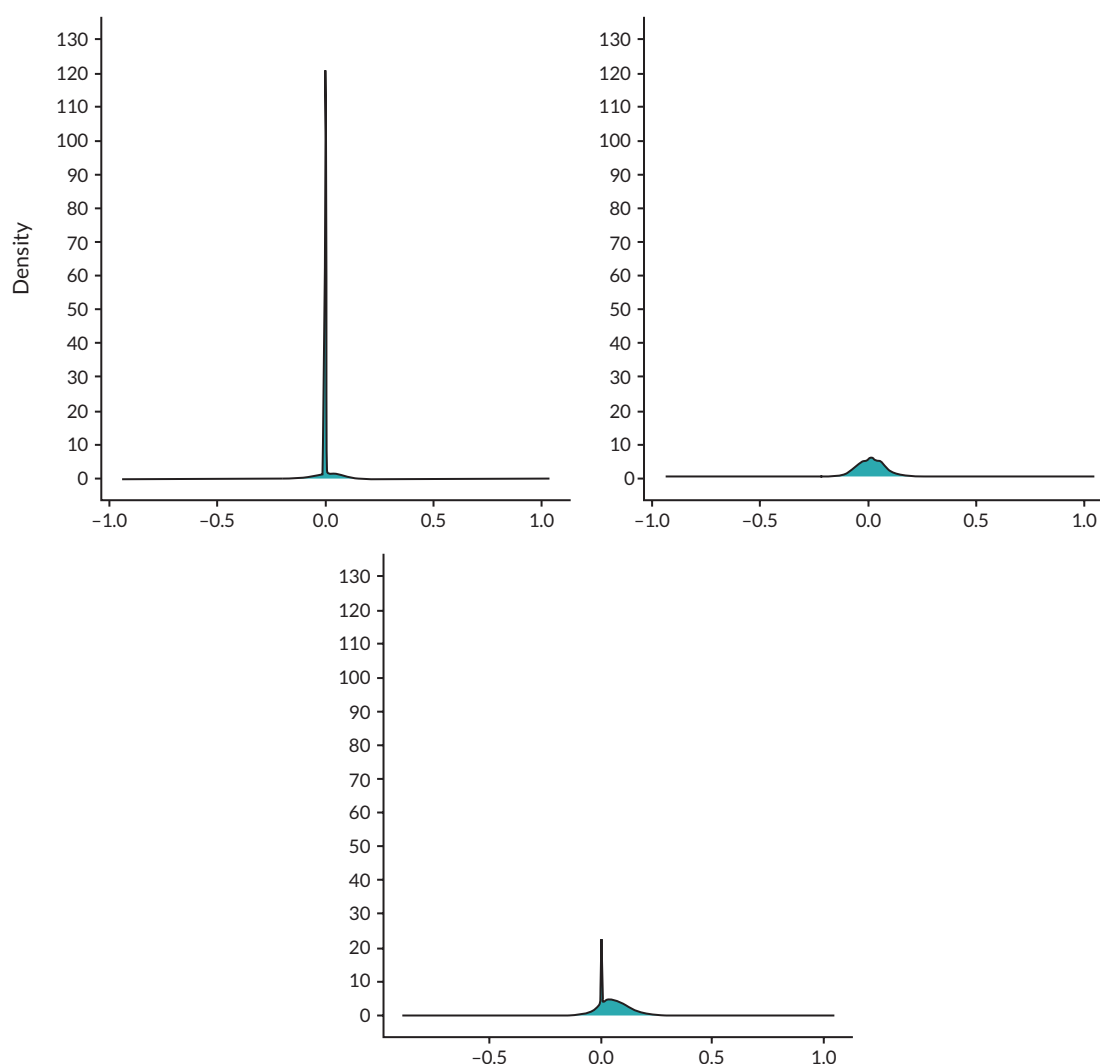
FIGURE 21 Distribution of change in mean EVI from left to right compared with previous quarter, previous year, and baseline.

shows a normal distribution centred around the no-change value (mean 0.01), with the majority of those in the cohort experiencing change in EVI within 1 SD (±0.12; *Figure 21*, central panel). Change in EVI compared with baseline (defined as the first period an individual enters the cohort) shows a slight skew towards positive change (mean change 0.05; *Figure 21*, right-hand side). This represents the 'greening' of Wales over the study period (2008–19) described in *Chapter 3*, where mean EVI increased by 0.13 in rural areas and 0.09 in urban areas.

Change in access to green and blue spaces among those who moved home only

The secondary environmental exposure was change in the number of potentially accessible GBS. In this study, only those people who moved home (42.5% of the cohort; *Table 2*) had the potential to experience a change in their access to GBS (see *Chapter 4*). People who moved home more than once experienced, on average, an increase in the number of potentially accessible GBS compared with baseline (change in mean number of GBS 0.65), while those who moved home only once experienced a decrease in the number of potentially accessible GBS (change in mean number of GBS −7.48; *Table 14*). Of all home moves, 31% resulted in an increase in GBS access and 31% resulted in a decrease (and 38% remained the same). Similarly, 35% resulted in an increase in EVI around the home and 33% in a decrease in EVI (*Appendix 5*, *Table 36*).

TABLE 13 Change in mean EVI and seeking help for CMD in general practice

	Mean EVI within 300 m of the home								
	Change compared with previous quarter			Change compared with previous year			Change compared with baseline		
Individuals (observations)	2,801,483 (91,746,917)			2,801,483 (83,244,259)			2,801,483 (91,739,536)		
	AOR	95% CI		AOR	95% CI		AOR	95% CI	
	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Time period within study	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01
Age (centred to mean)	1.01	1.01	1.01	1.00	1.00	1.01	1.01	1.01	1.01
Sex									
Male (ref)	1.00	-	-	1.00	-	-	1.00	-	-
Female	2.09	2.08	2.11	2.12	2.11	2.13	2.12	2.11	2.13
WIMD									
1 – Most deprived	1.18	1.17	1.18	1.20	1.19	1.20	1.18	1.17	1.18
2	1.12	1.11	1.12	1.13	1.12	1.13	1.11	1.11	1.12
3	1.08	1.08	1.08	1.08	1.07	1.08	1.07	1.07	1.08
4	1.04	1.03	1.04	1.03	1.03	1.03	1.03	1.02	1.03
5 – Least deprived (ref)	1.00	-	-	1.00	-	-	1.00	-	-
Moved home (in a given quarter)									
No (ref)	1.00	-	-	1.00	-	-	1.00	-	-
Yes	1.18	1.18	1.19	1.18	1.18	1.19	1.20	1.19	1.20
Birth in household									
No (ref)	1.00	-	-	1.00	-	-	1.00	-	-
Yes	0.84	0.83	0.85	0.83	0.82	0.84	0.85	0.84	0.85
Death in household									
No (ref)	1.00	-	-	1.00	-	-	1.00	-	-
Yes	1.19	1.17	1.20	1.18	1.17	1.20	1.19	1.17	1.20
Charlson Comorbidity Index	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01
Random Effects (I) Coefficient (SE)	5.65	0.01		5.64	0.01		5.69	0.01	

Results reported for a 0.1-unit change in mean EVI. I, individual level; SE, standard error.

Research question 1: Is greater residential exposure and access to green and blue spaces associated with the likelihood of seeking help for a common mental health disorder in general practice?

Change in enhanced vegetation index around the home and common mental health disorders

An increase of 0.1 EVI around the home did not affect seeking help for CMD in general practice, either before or after adjusting for covariates (*change in mean EVI relative to* previous quarter, year or baseline AOR 1.00, 95% CI 1.00 to 1.00; *Table 13*). Among the 1,611,581 people who did not move home

TABLE 14 Change in number of potentially accessible GBS and seeking help for CMD in general practice

Number of GBS within 1600 m of the home

Individuals (observations)	Change compared with previous quarter						Change compared with previous year						Change compared with baseline					
	Moved home once 622,025 (17,946,687)			Moved home > once 567,877 (17,026,204)			Moved home once 622,025 (16,189,423)			Moved home > once 567,877 (15,047,243)			Moved home once 622,025 (16,828,245)			Moved home > once 567,877 (15,345,319)		
	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median
	-0.22	46.61	0	-0.68	88.35	0	-0.83	91.32	0	-2.71	171.25	0	-7.48	231.98	0	0.65	306.35	0
	AOR	95% CI		AOR	95% CI		AOR	95% CI		AOR	95% CI		AOR	95% CI		AOR	95% CI	
Time period within study	1.01	1.01	1.01	1.02	1.02	1.02	1.01	1.01	1.01	1.02	1.02	1.02	1.01	1.01	1.01	1.02	1.02	1.02
Age (centred to mean)	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01
Sex																		
Male (ref)	1.00	-	-	1.00	-	-	1.00	-	-	1.00	-	-	1.00	-	-	1.00	-	-
Female	2.06	2.04	2.09	1.71	1.69	1.73	2.10	2.07	2.13	1.74	1.72	1.76	2.08	2.05	2.11	1.76	1.74	1.78
WIMD																		
1 – Most deprived	1.17	1.16	1.18	1.13	1.13	1.14	1.19	1.17	1.20	1.14	1.14	1.15	1.15	1.14	1.16	1.12	1.11	1.13
2	1.10	1.09	1.11	1.09	1.08	1.09	1.11	1.10	1.12	1.09	1.09	1.10	1.09	1.08	1.10	1.08	1.07	1.08
3	1.06	1.05	1.07	1.05	1.05	1.06	1.07	1.06	1.08	1.06	1.05	1.07	1.06	1.05	1.07	1.05	1.04	1.06
4	1.02	1.01	1.03	1.01	1.00	1.02	1.02	1.01	1.03	1.01	1.00	1.02	1.02	1.01	1.03	1.01	1.00	1.02
5 – Least deprived (ref)	1.00	-	-	1.00	-	-	1.00	-	-	1.00	-	-	1.00	-	-	1.00	-	-
Moved home (in a given quarter)																		
No (ref)	1.00	-	-	1.00	-	-	1.00	-	-	1.00	-	-	1.00	-	-	1.00	-	-
Yes	1.15	1.14	1.16	1.19	1.18	1.19	1.15	1.14	1.16	1.19	1.18	1.20	1.15	1.14	1.16	1.20	1.20	1.21

60

Table 14 Change in number of potentially accessible GBS and seeking help for CMD in general practice *(continued)*

Number of GBS within 1600 m of the home

	Change compared with previous quarter						Change compared with previous year						Change compared with baseline					
	Moved home once			Moved home > once			Moved home once			Moved home > once			Moved home once			Moved home > once		
Individuals (observations)	622,025 (17,946,687)			567,877 (17,026,204)			622,025 (16,189,423)			567,877 (15,047,243)			622,025 (16,828,245)			567,877 (15,345,319)		
	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median
Birth in household																		
No (ref)	1.00	-	-	1.00	-	-	1.00	-	-	1.00	-	-	1.00	-	-	1.00	-	-
Yes	0.83	0.81	0.85	0.83	0.82	0.84	0.82	0.81	0.84	0.82	0.81	0.83	0.83	0.81	0.85	0.84	0.82	0.85
Death in household																		
No (ref)	1.00	-	-	1.00	-	-	1.00	-	-	1.00	-	-	1.00	-	-	1.00	-	-
Yes	1.18	1.15	1.21	1.14	1.11	1.17	1.17	1.14	1.21	1.14	1.11	1.17	1.18	1.15	1.21	1.14	1.11	1.17
Charlson Comorbidity Index	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01
Random Effects (I) Coefficient (SE)	5.60 (0.01)			3.58 (0.01)			5.64 (0.01)			3.61 (0.01)			5.54 (0.01)			3.50 (0.01)		

I, individual level; SE, standard error.

(57.5% of the cohort; *Table 2*) large changes in mean EVI (increases or decreases of ≥ 0.15 EVI) were not associated with seeking help for CMD in general practice (AOR for a large increase in EVI relative to baseline 1.00, 95% CI 0.99 to 1.00; AOR for a large decrease in EVI relative to baseline 1.00, 95% CI 1.00 to 1.01; *Table 17*).

The odds of women seeking help for CMD in general practice were twice as high as the odds of men seeking help for CMD in general practice, regardless of increases in EVI around the home (*Table 13*). The odds of seeking help for CMD in general practice were higher with each quintile of increasing deprivation regardless of increase in EVI; those living in the most deprived areas had 20% higher odds of seeking help for CMD compared with those living in the least deprived areas (*Table 13*).

Change in access to green and blue spaces and common mental health disorders among those who moved home only

Among the 1,189,902 people who moved home at least once (42.5% of the cohort; *Table 2*), an increase in the number of potentially accessible GBS did not affect seeking help for CMD in general practice, either before or after adjusting for covariates (*change in total number of GBS relative t*o each of previous quarter, year and baseline for those who moved once and who moved more than once AOR 1.00, 95% CI 1.00 to 1.00; *Table 14*). Similar to the results for change in EVI and seeking help for CMD above, the odds of women seeking help for CMD in general practice were twice as high as the odds for men, regardless of increase in potentially accessible GBS. With each quintile of increasing deprivation, the odds of seeking help for CMD in general practice were higher.

Research question 2.1: Is the association between EVI exposure and seeking help for CMD in general practice modified by socioeconomic disadvantage?
Increases in EVI around the home did not significantly affect seeking help for CMD in general practice for those living in the most deprived urban areas, or those living in the least deprived urban areas (*Table 15*).

The odds of seeking help for CMD in general practice were higher in the quarter when an individual moved home, regardless of whether that move resulted in an increase in EVI (AOR 1.15, 95% CI 1.14 to 1.16 for those who moved once, AORs 1.19 to 1.20 for those who moved more than once, *Table 16*). We found the same pattern of higher odds of seeking help for CMD in general practice in the quarter an individual moved, regardless of whether the move resulted in an increase in the number of potentially accessible GBS (AOR 1.15, 95% CI 1.14–1.16 for those who moved once, AORs 1.19 to 1.20 for those who moved more than once, *Table 14*).

Research question 2.2: Is the association between EVI and seeking help for CMD in general practice modified by moving home?
Increases in EVI around the home affected seeking help for CMD in general practice for those who moved home more than once, with an increase of 0.1 EVI around the home associated with slightly lower odds of seeking help for CMD in general practice (*change in mean EVI relative to baseline*: AOR 0.99, 95% CI 0.99 to 0.99, *Table 16*). An increase in EVI (+ 0.1 EVI) around the home did not significantly affect the likelihood of seeking help for CMD in general practice (*Table 17*).

Research question 3: Do people with different GBS exposures through time have different associated risks of using general practice services?

Change in enhanced vegetation index around the home and general practice events
Our fitted model implies that an increase in EVI around the home (compared to the previous quarter, year and to baseline) did not significantly affect the number of GPE days in a quarter, either before or after adjusting for covariates. Adjusted incidence rate ratios (AIRR) for GPE days for a 0.1-unit increase in mean EVI relative to baseline 1.00, 95% CI 1.00 to 1.00 (*Table 18*). Women had a 1.4 times higher rate of GPE days than men, regardless of an increase in EVI, or other variables (see *Table 18*). The rate of GPE days was

TABLE 15 Change in mean EVI and seeking help for CMD among those only living in urban areas, by deprivation

Change in mean EVI within 300 m of the home

Individuals (observations)	Compared to previous quarter					Compared to previous year					Compared to baseline				
	All urban	Least deprived'		Most deprived'		All urban	Least deprived'		Most deprived'		All urban	Least deprived'		Most deprived'	
	1,782,459 (56,696,324)	278,885 (8,701,437)		336,889 (10,645,773)		1,782,459 (48,417,095)	278,885 (7,883,824)		336,889 (9,661,366)		1,782,459 (56,869,381)	278,885 (8,839,859)		336,889 (10,645,773)	
	AOR	AOR	95% CI	AOR	95% CI	AOR	AOR	95% CI	AOR	95% CI	AOR	AOR	95% CI	AOR	95% CI
	1.00	1.00	0.99 to 1.00	1.00	1.00 to 1.01	1.00	1.00	0.99 to 1.00	1.00	1.00 to 1.00	1.00	1.00	1.00 to 1.02	1.00	1.00 to 1.01
Time period within study	1.01	1.01	1.01 to 1.01	1.01	1.01 to 1.01	1.01	1.01	1.01 to 1.01	1.01	1.01 to 1.01	1.01	1.01	1.01 to 1.01	1.01	1.01 to 1.01
Age (centred to mean)	1.01	1.01	1.01 to 1.01	1.01	1.01 to 1.01	1.01	1.01	1.01 to 1.01	1.01	1.01 to 1.01	1.01	1.01	1.01 to 1.01	1.01	1.01 to 1.01
Sex															
Male (ref)	1.00	1.00	–	1.00	–	1.00	1.00	–	1.00	–	1.00	1.00	–	1.00	–
Female	2.10	2.18	2.13 to 2.23	2.19	2.15 to 2.23	2.14	2.18	2.13 to 2.23	2.23	2.19 to 2.26	2.13	2.15	2.10 to 2.20	2.20	2.16 to 2.23
WIMD															
1 – Most deprived	1.21	–	–	–	–	1.22	–	–	–	–	1.20	–	–	–	–
2	1.15	–	–	–	–	1.16	–	–	–	–	1.15	–	–	–	–
3	1.11	–	–	–	–	1.11	–	–	–	–	1.11	–	–	–	–
4	1.06	–	–	–	–	1.06	–	–	–	–	1.06	–	–	–	–
5 – Least deprived (ref)	1.00	–	–	–	–	1.00	–	–	–	–	1.00	–	–	–	–

Table 15 Change in mean EVI and seeking help for CMD among those only living in urban areas, by deprivation *(continued)*

	Change in mean EVI within 300 m of the home														
	Compared to previous quarter					Compared to previous year					Compared to baseline				
	All urban	Least deprived`		Most deprived`		All urban	Least deprived`		Most deprived`		All urban	Least deprived`		Most deprived`	
Individuals (observations)	1,782,459 (56,696,324)	278,885 (8,701,437)		336,889 (10,645,773)		1,782,459 (48,417,095)	278,885 (7,883,824)		336,889 (9,661,366)		1,782,459 (56,869,381)	278,885 (8,839,859)		336,889 (10,645,773)	
	AOR	AOR	95% CI	AOR	95% CI	AOR	AOR	95% CI	AOR	95% CI	AOR	AOR	95% CI	AOR	95% CI
Moved home (in a quarter)															
No (ref)	1.00	1.00	- -	1.00	- -	1.00	1.00	- -	1.00	- -	1.00	1.00	- -	1.00	- -
Yes	1.18	1.10	1.06 1.14	1.19	1.17 1.22	1.18	1.10	1.06 1.14	1.20	1.18 1.22	1.19	1.11	1.08 1.15	1.21	1.19 1.23
Birth in household															
No (ref)	1.00	1.00	- -	1.00	- -	1.00	1.00	- -	1.00	- -	1.00	1.00	- -	1.00	- -
Yes	0.85	0.86	0.81 0.91	0.87	0.85 0.89	0.84	0.86	0.81 0.91	0.86	0.84 0.89	0.85	0.84	0.80 0.89	0.87	0.85 0.90
Death in household															
No (ref)	1.00	1.00	- -	1.00	- -	1.00	1.00	- -	1.00	- -	1.00	1.00	- -	1.00	- -
Yes	1.19	1.20	1.14 1.26	1.22	1.18 1.25	1.19	1.20	1.14 1.25	1.21	1.17 1.25	1.19	1.21	1.16 1.27	1.21	1.18 1.25
Charlson Comorbidity Index	1.01	1.01	1.01 1.01	1.01	1.01 1.01	1.01	1.01	1.01 1.01	1.01	1.01 1.01	1.01	1.01	1.01 1.01	1.01	1.01 1.01
Random Effects (I) Coef. (SE*)	5.58 (0.01)	7.78 (0.01)		5.13 (0.01)		5.56 (0.01)	7.75 (0.01)		5.08 (0.01)		5.63 (0.01)	7.83 (0.01)		5.18 (0.01)	

Results reported for a 0.1 unit change in mean EVI. I, individual level; SE = standard error.

TABLE 16 Change in mean EVI and seeking help for CMD, by moving home

	Change in mean EVI within 300 m of the home																	
	Compared to previous quarter						**Compared to previous year**						**Compared to baseline**					
	Did not move home		Moved home once		Moved > once		Did not move home		Moved home once		Moved > once		Did not move home		Moved home once		Moved > once	
Individuals (observations)	1,611,581 (53,118,438)		622,025 (19,771,721)		567,877 (18,856,758)		1,611,581 (48,417,095)		622,025 (17,923,914)		567,877 (16,903,250)		1,611,581 (54,557,404)		622,025 (19,253,861)		567,877 (17,928,271)	
	AOR	95% CI	AOR	95% CI	AOR	95% CI	AOR	95% CI	AOR	95% CI	AOR	95% CI	AOR	95% CI	AOR	95% CI	AOR	95% CI
Time period within study	1.01	1.01 1.01	1.01	1.01 1.01	1.02	1.02 1.02	1.01	1.01 1.01	1.01	1.01 1.01	1.02	1.02 1.02	1.01	1.01 1.01	1.01	1.01 1.01	0.99	0.99 0.99
Age (centred to mean)	1.01	1.01 1.01	1.01	1.01 1.01	1.01	1.01 1.01	1.01	1.01 1.01	1.01	1.01 1.01	1.01	1.01 1.01	1.01	1.01 1.01	1.01	1.01 1.01	1.02	1.02 1.02
Sex																		
Male (ref)	1.00	–	1.00	–	1.00	–	1.00	–	1.00	–	1.00	–	1.00	–	1.00	–	1.00	–
Female	2.25	2.24 2.27	2.06	2.03 2.08	1.71	1.69 1.73	2.29	2.27 2.31	2.09	2.07 2.12	1.73	1.71 1.75	2.25	2.23 2.27	2.07	2.05 2.10	1.75	1.73 1.77
WIMD																		
1 – Most deprived	1.40	1.39 1.42	1.16	1.15 1.17	1.12	1.11 1.13	1.41	1.40 1.43	1.17	1.16 1.18	1.12	1.11 1.13	1.40	1.38 1.41	1.14	1.13 1.15	1.11	1.10 1.12
2	1.25	1.24 1.26	1.09	1.08 1.10	1.07	1.07 1.08	1.26	1.25 1.27	1.10	1.09 1.11	1.08	1.07 1.09	1.25	1.24 1.26	1.08	1.07 1.09	1.07	1.06 1.07
3	1.14	1.13 1.15	1.06	1.05 1.06	1.05	1.04 1.05	1.14	1.13 1.15	1.06	1.05 1.07	1.05	1.04 1.06	1.14	1.13 1.15	1.05	1.04 1.06	1.05	1.04 1.05
4	1.06	1.05 1.07	1.02	1.01 1.02	1.01	1.00 1.01	1.06	1.05 1.07	1.02	1.01 1.03	1.01	1.00 1.01	1.06	1.05 1.07	1.02	1.01 1.02	1.02	1.00 1.02
5 – Least deprived (ref)	1.00	–	1.00	–	1.00	–	1.00	–	1.00	–	1.00	–	1.00	–	1.00	–	1.00	–
Moved home (in a given quarter)																		
No (ref)	–		1.00	–	1.00	–	–		1.00	–	1.00	–	–		1.00	–	1.00	–
Yes	–		1.15	1.14 1.16	1.19	1.18 1.19	–		1.15	1.14 1.16	1.19	1.18 1.19	–		1.15	1.14 1.16	1.20	1.20 1.21

continued

Table 16 Change in mean EVI and seeking help for CMD, by moving home *(continued)*

	Change in mean EVI within 300 m of the home								
	Compared to previous quarter			Compared to previous year			Compared to baseline		
	Did not move home	Moved home once	Moved > once	Did not move home	Moved home once	Moved > once	Did not move home	Moved home once	Moved > once
Individuals (observations)	**1,611,581 (53,118,438)**	**622,025 (19,771,721)**	**567,877 (18,856,758)**	**1,611,581 (48,417,095)**	**622,025 (17,923,914)**	**567,877 (16,903,250)**	**1,611,581 (54,557,404)**	**622,025 (19,253,861)**	**567,877 (17,928,271)**
	AOR 95% CI	AOR 95% CI	AOR 95% CI	AOR 95% CI	AOR 95% CI	AOR 95% CI	AOR 95% CI	AOR 95% CI	AOR 95% CI
Birth in household									
No (ref)	1.00 - -	1.00 - -	1.00 - -	1.00 - -	1.00 - -	1.00 - -	1.00 - -	1.00 - -	1.00 - -
Yes	0.87 0.85 0.89	0.83 0.81 0.85	0.93 0.82 0.84	0.88 0.86 0.90	0.83 0.81 0.84	0.80 0.83	0.87 0.85 0.89	0.83 0.82 0.85	0.83 0.82 0.84
Death in household									
No (ref)	1.00 - -	1.00 - -	1.00 - -	1.00 - -	1.00 - -	1.00 - -	1.00 - -	1.00 - -	1.00 - -
Yes	1.21 1.19 1.23	1.17 1.14 1.20	1.13 1.10 1.16	1.21 1.19 1.23	1.16 1.14 1.19	1.13 1.10 1.16	1.21 1.19 1.23	1.17 1.14 1.20	1.13 1.10 1.16
Charlson Comorbidity Index	1.01 1.01 1.01	1.01 1.01 1.01	1.01 1.01 1.01	1.01 1.01 1.01	1.01 1.01 1.01	1.01 1.01 1.01	1.01 1.01 1.01	1.01 1.01 1.01	1.01 1.01 1.01
Random Effects (I) Coef. (SE*)	7.04 (0.01)	5.57 (0.01)	3.56 (0.01)	7.00 (0.01)	5.64 (0.01)	3.58 (0.01)	7.07 (0.01)	5.50 (0.01)	3.47 (0.01)

Results reported for a 0.1 unit change in mean EVI. I, individual level; SE, standard error.

TABLE 17 Change in mean EVI (categorical) and seeking help for CMD among those who do not move home

	Change in mean EVI within 300 m of the home					
	Compared to previous year			Compared to baseline		
Individuals (observations)	1,611,581 (48,417,095)			1,611,581 (54,557,404)		
	AOR	**95% CI**		**AOR**	**95% CI**	
Large negative change in mean EVI	1.00	1.00	1.00	1.00	1.00	1.01
No/small change (ref)	1.00	-	-	1.00	-	-
Large positive change in mean EVI	1.00	1.00	1.01	1.00	1.00	1.01
Time period within study	1.01	1.01	1.01	1.01	1.01	1.01
Age (centred to mean)	1.01	1.01	1.01	1.01	1.01	1.01
Sex						
Male (ref)	1.00	-	-	1.00	-	-
Female	2.29	2.27	2.31	2.25	2.23	2.27
WIMD						
1 – Most deprived	1.41	1.40	1.43	1.40	1.38	1.41
2	1.26	1.25	1.27	1.25	1.24	1.26
3	1.14	1.13	1.15	1.14	1.13	1.15
4	1.06	1.05	1.07	1.06	1.05	1.07
5 – Least deprived (ref)	1.00	-	-	1.00	-	-
Birth in household						
No (ref)	1.00	-	-	1.00	-	-
Yes	0.88	0.86	0.90	0.87	0.85	0.89
Death in household						
No (ref)	1.00	-	-	1.00	-	-
Yes	1.21	1.19	1.23	1.21	1.19	1.23
Charlson Comorbidity Index	1.01	1.01	1.01	1.01	1.01	1.01
Random Effects (I) Coef. (SE*)	7.00 (0.01)			7.07 (0.01)		

I, individual level; SE, standard error.

higher with each quintile of increasing deprivation, although the magnitude of differences was small (AIRR for those living in the most deprived areas compared to those living in the least deprived areas: 1.03, 95% CI 1.03 to 1.03, *Table 18*).

Change in access to green and blue spaces and use of general practice services among those who moved home only

Our modelling suggests that an increase in the number of potentially accessible GBS did not significantly affect general practice service utilisation, measured by the number of GPE days in a quarter (*Table 19*), before or after adjusting for covariates (*change in total GBS count relative to*: previous quarter AIRR 1.00, 95% CI 1.00 to 1.00; previous year AIRR 1.00, 95% CI 0.99 to 1.00; baseline AIRR 1.00, 95% CI 1.00 to 1.00).

TABLE 18 Change in mean EVI change and number of GPE days in a quarter

	Change in mean EVI within 300 m of the home									
	Compared with previous quarter				Compared with previous year			Compared with baseline		
Individuals (observations)	2,801,483	2,801,483 (91,746,917)			2,801,483 (83,244,259)			2,801,483 (91,739,536)		
	IRR	AIRR	95% CI		AIRR	95% CI		AIRR	95% CI	
Exposure change	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Time period within study	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Age (centred to mean)	1.04	1.03	1.03	1.03	1.03	1.03	1.03	1.03	1.03	1.03
Sex										
Male (ref)	1.00	1.00	-	-	1.00	-	-	1.00	-	-
Female	1.30	1.42	1.42	1.42	1.43	1.42	1.43	1.41	1.41	1.41
WIMD										
1 – Most deprived	0.99	1.03	1.03	1.03	1.03	1.03	1.03	1.03	1.03	1.03
2	1.01	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02
3	0.99	1.01	1.01	1.02	1.01	1.01	1.02	1.01	1.01	1.02
4	0.99	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01
5 – Least deprived (ref)	1.00	1.00	-	-	1.00	-	-	1.00	-	-
Moved home (in a given quarter)										
No (ref)	1.00	1.00	-	-	1.00	-	-	1.00	-	-
Yes	1.51	1.49	1.49	1.49	1.50	1.50	1.50	1.50	1.50	1.51
Birth in household										
No (ref)	1.00	1.00	-	-	1.00	-	-	1.00	-	-
Yes	1.43	1.46	1.49	1.49	1.45	1.44	1.45	1.45	1.44	1.45
Death in household										
No (ref)	1.00	1.00	-	-	1.00	-	-	1.00	-	-
Yes	1.25	1.02	1.01	1.02	1.02	1.01	1.02	1.01	1.01	1.02
Charlson Comorbidity Index	1.03	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02
Random Effects (I) Coef. (SE)		0.81 (0.00)			0.82 (0.00)			0.76 (0.00)		

Results reported for a 0.1 unit change in mean EVI. I, individual level; SE, standard error.

Economic impact assessment of change in use of general practice services associated with change in enhanced vegetation index

In *Chapter 4*, we described an overall increase in mean EVI in Wales between 2008 and 2019 of approximately 0.1 EVI, likely related to natural 'greening' occurring as vegetation matures. Increases in EVI of this magnitude did not significantly affect general practice service utilisation in our study (measured by GPE days per quarter). We therefore could not calculate the potential impact of the proportional change in service utilisation on the Welsh general practice budget associated with an increase in EVI of 0.1.[196]

TABLE 19 Change in number of GBS and number of GPE days in a quarter

Individuals (observations)	Change in number of GBS within 1600 m of the home Individuals who move								
	Compared with previous quarter			Compared with previous year			Compared with baseline		
	1,189,902 (34,972,891)			1,189,902 (31,236,666)			1,189,902 (32,173,564)		
	IRR	AIRR	95% CI		AIRR	95% CI		AIRR	95% CI	
Exposure change	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Time period within study	1.00	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01
Age (centred to mean)	1.03	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02
Sex										
Male (ref)	1.00									
Female		1.49	1.49	1.50	1.50	1.50	1.50	1.49	1.49	1.50
WIMD										
1 – Most deprived		1.03	1.03	1.03	1.03	1.03	1.03	1.02	1.02	1.03
2		1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02
3		1.01	1.01	1.01	1.01	1.01	1.02	1.01	1.01	1.02
4		1.00	1.00	1.01	1.01	1.00	1.01	1.01	1.00	1.01
5 – Least deprived (ref)	1.00									
Moved home (in a given quarter)										
No (ref)	1.00									
Yes		1.50	1.50	1.50	1.50	1.50	1.50	1.51	1.51	1.51
Birth in RALF										
No (ref)	1.00									
Yes		1.51	1.51	1.52	1.50	1.50	1.51	1.51	1.50	1.51
Death in RALF										
No (ref)	1.00									
Yes		1.01	1.01	1.02	1.01	1.01	1.02	1.01	1.01	1.02
Charlson Comorbidity Index	1.03	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02
Random Effects (I) Coef. (SE)		0.82 (0.00)			0.83 (0.00)			0.76 (0.00)		

I, individual level; SE, standard error.

Discussion

We addressed the evidence gap on whether changes in the 'greenness and blueness' of environments affect future adult mental health risk. We summarise our answers to each of our research questions, then discuss these findings in context.

Findings by research question

Research question 1: Do people with different exposures to green and blue spaces through time have different associated risks of having a common mental health disorder?

In these longitudinal analyses, we considered differences in exposures and outcomes for the same person. We found that increases in GBS exposure, in terms of ambient greenness (EVI, whole cohort) or the number of potentially accessible GBS (among those who had moved home), did not affect the likelihood of subsequently seeking help for CMD in general practice. For those who did not move home, large increases or decreases in EVI did not affect seeking help for CMD in general practice.

Research question 2: Are changes in green and blue space exposures and seeking help for common mental health disorders in general practice modified by socioeconomic disadvantage and moving home?

We found no evidence that increases in GBS exposure over time and subsequent help seeking for CMD in general practice varied by socioeconomic disadvantage. Increases in EVI around the home affected seeking help for CMD in general practice for those who moved home more than once, (with very slightly lower odds of seeking help for CMD) but not those who only moved home once, or who did not move.

Moving home was associated with increases in seeking help for CMD in general practice in the quarter of the move. Overall, however, we found no association between increases in GBS exposures in the previous quarter or year and the likelihood of seeking help for CMD in general practice. Our null results may be explained by the following reasons. When an individual's initial mental health response to moving home (or their reasons for moving) were negative at the time of the move, their mental health may positively rebound after their move, related to an improved environment. This could be due to ambient greenness or GBS access, or other factors associated with the move, for example better accommodation, a shorter commute, living close to family or friends. This would render our results null due to the initial negative stressor of the move, offset by a regain in positive mental health once settled into a new community. Our analyses were structured to capture seeking help for CMD in each quarter following prior changes in GBS exposures. Moving home was flagged in the same quarter as seeking help for CMD was examined. In our stratified analyses to answer research question 2, we identified individuals who had moved home (once, or more than once) and examined associations between their changes in GBS exposure and subsequent CMD. However, changes in EVI exposure were not temporally linked to home move(s).

In situ changes in EVI could have occurred. For example, a change in the GBS itself such as building on a local park (reduction), or a new community garden on vacant ground (gain). Or changes could occur at the individual level, such as through moving home. Despite mitigations taken in data processing, seasonality could also have introduced changes in EVI (see *Limitations* below for more information). In our data, changes in the number of potentially accessible GBS could only occur when an individual moved home. For those who moved only once, we could attribute specific changes in the number of accessible GBS to a specific home move. In this group however, we found no evidence that an increase in the number of potentially accessible GBS significantly affected seeking help for CMD. As above, exposures were coincident with help seeking for CMDs in the same quarter as the move; we did not capture changes in future years.

In future work we plan to examine associations between changes in GBS exposures arising from moving home and seeking help for CMD both at the time of the move, and in future periods following the move. We will explore the hypothesis that initial negative impacts on mental health associated with moving home can be offset by subsequent improvements in mental health when the move is to a greener area. Previous studies have shown that mental health improved within a year of moving to a greener area, and that this was maintained for the following two years.[45]

Research question 3: Do people with different green and blue space exposures through time have different associated risks of using general practice services?

Over our study period, the mean number of GPE days per person per quarter increased by 1 day from 3.3 to 4.3 days. Given growing pressures on general practice service provision, even a small decrease in general practice service use at the individual level associated with increased GBS exposure may have a high value in releasing resources within these services. However, we found no evidence that increases in GBS exposure, in terms of ambient greenness (EVI) or the number of potentially accessible GBS, significantly affected general practice service utilisation (measured by the number of GPE days in a quarter).

Findings in context

Living close to GBS such as parks, woodlands, trails, ponds, lakes, rivers and beaches is associated with positive impacts on physical and mental health.[3,4,23,198] However, the majority of evidence has not unpicked associations between the type, proximity, quantity and 'qualities' of GBS, and changes in mental health/well-being.[30,73] As a result, existing evidence to inform policies shaping our environment is limited.[60,82] In *Chapter 5*, we found that those with overall higher levels of ambient greenness around the home, or an overall greater number of potentially accessible GBS throughout their time in the study had lower odds of seeking help for CMD in general practice at any point in the study.

Those time-aggregated results add to growing evidence suggesting that GBS have positive impacts on mental health.[1–3,139] However, there is only limited evidence for a causal relationship between the 'greenness and blueness' of environments and adult mental health,[8] with a relatively limited amount of robust, longitudinal evidence.[8,45,82] This chapter reports our national, longitudinal evaluation of the impact of changes in ambient greenness and availability of potentially accessible GBS around the home on seeking help for CMD in general practice among adults in Wales.

We found no evidence that changes in EVI or the number of potentially accessible GBS over time impacted on people's propensity to seek help for a CMD in general practice. We suggest these findings are indicative of the limitations of the data and modelling approach. In *Chapter 7*, we examine associations between leisure use of GBS with well-being, and seeking care for CMD in general practice, by linking data from the NSW to the cohort.

Strengths

This is the first e-cohort with national household-level longitudinal annual environment metrics linked to longitudinal EHRs. Our work has several strengths due to the available data and the analytical approach. First, many studies of the association between the environment and mental health have focused on 'greenness' (measured by EVI or NDVI). We have supplemented this with measures of GBS availability, based on a large, integrated, national-scale spatial dataset of potentially accessible spaces.

Many studies have used environmental data aggregated over small areas. By using the residential data linkage available in the SAIL databank, we have linked the environmental data for each home to the full cohort of 2.8 million people and their routinely collected, administrative health data. This enabled objective assessment of environmental changes, with no research burden for individuals.[104,108] Our approach reduces ecological fallacy by using accurate data for each home and allows us to anonymously follow home moves, capturing associated larger changes in GBS exposures.

The large adult population of the cohort provides sufficient power to examine variations between subgroups to investigate inequalities. The capability for anonymous follow-up also reduces the risk of unintentionally capturing 'green gentrification' effects when improvements to GBS are made in an area. This is a risk to area-level study designs; the original population may be replaced with a healthier population as an area gentrifies.[199] This is a particular issue over the longer timescale needed to evidence the impact of smaller, cumulative environmental changes. Place-based intervention studies investigating area-level health effects over long periods of time following environment improvements

risk recording health outcomes of a different population.[200] Our ability to stratify analyses according to whether individuals had moved reduced our exposure to these health selection effects.

Limitations

There are several limitations of the longitudinal analyses presented in this chapter. The primary limitation is that we observed high levels of ambient greenness across urban and rural areas, and smaller greener conurbations, in Wales. Households did not generally experience large changes in EVI values. This limited the amount of change in EVI observed across the period using Landsat remotely sensed imagery, impacting our ability to find associations between change in EVI and future chance of having a CMD.

Additionally, EVI is a measure of vegetation, which we used to estimate ambient greenness (as for many other health research studies).[64] EVI values below 0 represent water bodies and, for the most part, were excluded from our modelling by clipping imagery to coastal areas. While we clipped buffers in coastal areas to historical high-tide marks, our EVI metrics are still likely to have been impacted by the relatively high concentration of coastal and inland waters in and around Wales. Negative EVI values produced by surface water coupled with the potential inverse associations between blue space and CMD, mean that the lack of association observed between an increase in mean EVI and CMD may have been impacted by lower levels of CMD in 'bluer' areas. That said, the results for GBS accessibility did include blue spaces and were materially no different to those using EVI.

There are additional considerations when using change in EVI in longitudinal analyses. As described in *Chapter 3*, we were restricted to using images from Springtime (May–July) in Wales each year to maximise image quality because cloud cover. However, natural fluctuation in EVI with the seasons in temperate zones can result in changes as large at 0.2–0.3 EVI between winter and spring.[201] Combined with the three-month May to July window we applied to maximise image quality, this means that EVI may not have been captured at peak greenness every year. When EVI is used to measure change in ambient greenness in longitudinal studies, these issues may introduce noise to EVI data. This noise might exceed real variation in ambient greenness, resulting in potential exposure misclassification particularly when the observed changes are small. However, confidence in our results is strengthened by the comparable findings between the two GBS exposure measures (change in mean EVI and number of potentially accessible GBS).

The OSMM Topographic Layer used to create the GBS access measures is a rich dataset, but it was not designed to capture discrete longitudinal change in land use. During our study, we uncovered limitations such as changing land use classifications in the database being due to continual improvement programmes rather than real changes in land use (described in detail in *Chapter 3*). This meant we were unable to use these data to reliably measure temporal changes in GBS access. The results for change in GBS access and future chance of having a CMD were therefore restricted to those who moved home. This resulted in a smaller cohort (*n* = 622,025) compared with the cohort available for analyses of the impact of change in mean EVI (*n* = 2,801,483), our primary exposure. People who moved home may differ from those who did not move home in other ways that may affect the chance of having a CMD. This residential self-selection bias can result from unmeasured neighbourhood selection factors related to environmental exposures and CMD. Neighbourhood selection factors could include preference for greener areas or those with high-quality outdoor spaces, which could affect neighbourhood choice and CMD.[202]

Our analyses of GBS availability also do not distinguish between different types, sizes or qualities of spaces, and therefore we would expect heterogeneity in what this exposure measure reflects, and therefore the mechanisms through which these spaces might be expected to affect mental health.

The analytical approach employed here considers the relatively immediate impacts of a change in exposure. Previous evidence suggests that mental health benefits may arise during a longer time period

following a move to a greener area. These analyses may miss mental health impacts that take a longer time to arise following a change in GBS exposure.[45]

The validity and reliability of research using data from routinely collected EHRs depends upon its quality and completeness. Overall, the validity of primary care diagnoses in the UK tends to be high.[203] Case finding for CMD in routinely collected administrative health data can unobtrusively identify patients for mental health research, including on the effects of intervention.[204] Data derived from EHRs are predominantly routinely recorded and lack data on behaviour, some potential confounding factors, and outcomes such as well-being. For example, there is no health-related quality of life instrument routinely used to assess changes in health status in general practice in Wales. Because of this, the longitudinal analyses in this chapter are restricted to detecting changes in outcomes that involve health service utilisation. However, through linkage to NSW survey data, a subset of the cohort has information on well-being. We examine associations between GBS exposures and well-being in *Chapter 7*.

Conclusion

We combined EHRs with open-source satellite data to capture annual ambient exposure to greenness, temporally matched to subsequent health outcomes, improving on previous studies that did not have the data or systems to achieve this. We found no evidence that change in ambient greenness, or potential access to GBS around the home, affected either seeking help for CMD in general practice or use of general practice services. However, our work was limited by high levels of, and limited variation in, ambient greenness in Wales. There are specific limitations to these analyses that may affect results, and findings should be interpreted accordingly.

Chapter 7 Neighbourhood and leisure exposure to green and blue spaces and individual-level well-being and mental health

Introduction

Wellbeing is thought to be multidimensional, incorporating the balance of positive and negative emotions (affect); an evaluation of how one's life is going (evaluative); and purpose or feeling life is worthwhile (eudaimonic).[12,106,205] Promoting well-being has been recognised as a preventative public health strategy,[206] as high well-being is associated with a reduced risk of mental health disorders.[207] Depressive and anxiety disorders are the most prevalent mental health disorders[208] and, when grouped together, are termed 'common mental health disorders'.[209]

There is a substantial body of evidence that exposure to GBS is positively associated with psychological well-being and mental health.[31,41,61,82,210] There are many proposed pathways and mechanisms that may explain relationships between GBS exposure and well-being,[31] such as psychological restoration,[169] reduced stress and anxiety,[26,211] mitigation of urban noise,[7] and increased social cohesion.[41] Although some of these benefits could be attained by living near GBSs (e.g. restoration through visual exposure, noise mitigation[61]), other benefits (e.g. physical activity, social interactions) are assumed to be a direct result of using these local GBSs for recreation. For example, time spent walking was found to partly explain the relationship between coastal proximity and both mental health and general health in an English study.[43]

It seems intuitive that some benefits are dependent on actual usage of GBS, but few studies have investigated direct evidence for this assumption. There is fairly robust evidence that intentional recreational use of GBS is associated with subjective well-being and good mental health.[30,62,63] Although this evidence is primarily cross-sectional, allowing for the possibility of reverse causality (i.e. people with better well-being and mental health visit GBS more frequently), recent evidence suggests that individuals with CMDs are equally or more likely to make visits than those without.[212] It remains unclear to what extent people use nearby GBS versus travelling further afield, and this is likely to vary according to individual characteristics and circumstances.[68,213–216]

Studies considering mediation of the relationship between green space and health outcomes were recently reviewed.[64] Physical activity was the most common mediator considered, followed by social interaction and air pollution. Among the few studies that have explored the potential for mediation of neighbourhood GBS effects by visits, results are mixed. For example, across four European cities ($n = 3765$), it was estimated that time spent visiting nature close to home partially mediated the relationship between surrounding greenness (NDVI within 300 m of home) and mental health. However, this varied across cities and evidence for mediation was not found for the UK city.[217] Another study found evidence of mediation for the relationship between public park coverage (within 800 m) and mental health by park use in Mexico, but not for park coverage within 400 m.[218]

Physical activity is a mediator of particular interest as it is considered one of the key mechanisms through which exposure to GBS leads to mental and physical health benefit.[61] For example, using the nearest park for physical activity was found to mediate the relationship between park proximity and mental health in New York, although only for those not concerned with park safety.[219] Land-based physical activity (primarily walking) has been found to mediate the relationship between coastal proximity and both general health and mental health in England.[43] Walking is also an activity that has relatively few barriers to participation compared with other activities that may require specific equipment. Further, walking is the most common activity that takes place on recreational visits to natural environments.[140]

A number of factors appear to modify relationships between GBS exposure and mental health and well-being, including deprivation and type of GBS.[3,44,62] Effect modification by socioeconomic status is of particular interest, as it is a notable modifier of other health and GBS relationships[69,73,174,220,221] and aligns with a wide range of policy interests relating to health inequalities.[222-224] The extent of effect modification by socioeconomic status on relationships between leisure visits to GBS and well-being is not understood, but may be important.

The implications of demonstrating the relative importance of local GBS for well-being for urban and land use planning are clear. In particular, there is a need to better understand where to prioritise investment for equitable availability of GBS or for the targeted delivery of interventions in deprived communities to reduce health inequalities.[225]

Objectives

The overall aim of this chapter is to complement investigations described in earlier chapters. It does so by additionally considering: (1) exposure to GBS through intentional leisure use; (2) subjective well-being as an outcome; (3) potential mediation by physical activity (specifically walking); and (4) effect modification by urban/rural setting and socioeconomic deprivation. The SAIL cohort data do not contain data on intentional leisure visits or subjective well-being, but we were able to investigate these questions using NSW data linked to the SAIL cohort.

We examined the following research questions:

1. Is greater residential GBS exposure associated with higher subjective well-being, and is this association mediated by intentional GBS use for leisure?
2. Is greater residential GBS exposure associated with lower CMD likelihood, and is this association mediated by intentional GBS use for leisure?
3. Are associations between residential GBS exposure and use, and subjective well-being mediated by walking-specific physical activity?
4. Are associations between residential GBS exposure and use, and subjective well-being modified by individual-level deprivation?

Question 3 deviates slightly from the study protocol, which specified testing for mediation by total physical activity ('level of engagement in 150 minutes or more of moderate or vigorous intensity activity per week') and 'Proportion of all physical activity which is undertaken in green or blue space'. Data availability meant the latter mediation could not be tested because we could not identify whether physical activity was undertaken in GBS or elsewhere, such as gyms or swimming pools. Consequently, mediation analysis was conducted using weekly duration of moderate-vigorous physical activity gained through walking, given that most active visits to GBS involve walking,[89] and within the data available for this analysis, it was the most appropriate physical activity measure.

Methods

Data

National Survey for Wales data
The NSW (http://gov.wales/national-survey-wales) is an omnibus survey commissioned by the Welsh Government and carried out by the ONS. A NRW module of the NSW included questions on environmental issues and use of outdoor spaces. For those who consented, NSW responses from individuals were linked into the SAIL databank to their routinely collected health data. Data were derived from responses to the 2016/17 and 2018/19 surveys.[226,227] Data from 2017/18 were not used because of differences in the NRW survey module in that year, with relevant questions not asked.

Full details of survey methodologies can be found in official NSW technical reports,[102,228] but a brief overview is given here. The NSW interviews were conducted in person at participants' homes by trained ONS interviewers, using portable computers and showcards. Eligible interviewees were aged 16 years and over and drawn from a sample of randomly selected residential addresses in Wales, with one participant per residential address allowed. The sample of addresses was stratified by local authority (*n* = 22) and survey effort was approximately proportional to local authority populations, with the intent to obtain a nationally representative sample. Surveys took place throughout the year, so include responses across seasons. Certain modules of the questionnaire were only asked of a random subsample of total participants, including the NRW module and physical activity questions. Survey weights were provided that compensate for differences in sampling probability between different types of household and for differences between subsample and population profiles; analysis conducted by the NSW indicated that non-response bias was negligible and weights therefore do not adjust for this.

Outcomes

The primary outcome was a measure of subjective well-being: the WEMWBS.[106] The two secondary outcomes were subjective well-being as measured by life satisfaction[107] and CMDs (comprising symptoms, diagnoses and treatments for anxiety and depressive disorders, *Chapter 2*), identified using linked WLGP data (*Chapter 2*).

Warwick–Edinburgh Mental Well-being Scale

WEMWBS is a metric intended to capture affective, evaluative and eudaimonic well-being dimensions.[106] It is often applied in studies investigating relationships between well-being and GBS.[56,82,229] Participants are presented with a series of 14 positively worded statements regarding their thoughts or feelings, such as 'I've been feeling good about myself' and 'I've been thinking clearly'. They are asked to indicate how often they have had these thoughts or feelings within the last two weeks on a five-point scale: 'none of the time', 'rarely', 'some of the time', 'often', 'all of the time'. These responses are scored from one ('none of the time') to five ('all of the time') and summed, giving an overall score between 14 and 70 (higher scores indicate more positive mental well-being).

WEMWBS scores are designed to assess population level mental well-being.[106] They are typically normally distributed, with only minor (if any) floor and ceiling effects and are sensitive to changes in populations with both high and low baselines.[106,230] WEMWBS may be sensitive at the individual level,[230] but it is not intended to be a diagnostic measure. There are no established thresholds to indicate 'good' or 'bad' mental well-being,[231] though thresholds indicating 'low' or 'high' well-being are sometimes applied (e.g. lowermost and uppermost 15%).[229] We treated WEMWBS as a continuous variable.

Life satisfaction

Life satisfaction is a measure of evaluative well-being based on responses to the question, 'Overall, how satisfied are you with your life nowadays?'[107] Responses are scored zero 'not at all satisfied' to 10 'completely satisfied'. This gives a longer-term view of an individual's mental well-being than WEMWBS, and has also often been used in studies investigating well-being and GBS.[37,56,62,63] Some studies dichotomise scores into 'low' (0–7) and 'high' (8–10) scores based on distributions about median values,[63,156] but here they were treated as a continuous variable.

Common mental health disorders

NSW participants who provided consent were linked in SAIL using ALFs as described in earlier chapters. The likelihood of an individual NSW participant having a CMD was determined using the same algorithm described in *Chapter 2*: that is, by applying criteria from[103] to produce a binary outcome (high likelihood; low likelihood). NSW participants with a CMD flag for any quarter of the survey year were classified as

having a CMD. For ease of discussion, individuals are subsequently described as people with or without a CMD (as defined by the algorithm).

Exposure to green and blue spaces

Two measures of residential GBS (ambient greenness and access/proximity), and one of use (leisure time spent outdoors visiting GBS) were considered as exposures. These are described below and summarised in *Appendix 6*, *Table 37*.

Ambient greenness and proximity to green and blue spaces

GBS access/proximity and ambient greenness measures described in earlier chapters were available for NSW participants via linkage in SAIL. Full details are provided in *Chapter 3* but, in summary, EVI (yearly mean) provides an estimate of the relative ambient greenness within a 300 m buffer around each participant's home, theoretically recording values from –1 to +1, with healthy vegetation values found in the 0.2 to 0.8 range. Proximity to GBS indicates the road/footpath network distance to the nearest potentially accessible GBS, capped at 1600 m.

GBS use/visits

To estimate GBS use/visits, data were analysed for NSW participants based on questions about their visits to the outdoors in Wales. The survey specifies these visits as follows:

> The next questions are about outdoor recreation in Wales. We are interested in leisure visits and excursions to the Welsh outdoors of any length. These visits may have been made from your home or during holidays. By outdoors, we mean open spaces anywhere in the countryside or in towns and cities, including your local neighbourhood, paths, woodland, parks and farmland. Visits may have involved both active and passive pursuits.

Time spent on these leisure visits is subsequently referred to here as 'time outdoors'. Questions used to derive time outdoors from the NSW included those establishing which activities had been undertaken during these visits, and how long was spent doing them.[226,227] Weekly time outdoors was estimated as described in *Appendix 6*, *Table 37*. Owingto heavy skewing in the distribution, this value was capped at 420 minutes per week (i.e. values greater than 420 were re-coded as 420). This is comparable to the approach previously taken with English data for similar analyses.[63] Fifteen percent of values were recoded this way.

Covariates

Variables to be included in statistical models were selected based on data availability and theoretical understanding of their potential importance and associations with either/both exposure and outcome variables. A directed acyclic graph (*Appendix 6*, *Figures 43* and *44*) was developed based on the published literature as an illustrative tool to summarise potential relationships, as well as aiding identification of potential confounders, mediators and effect modifiers. For example, we identified that material deprivation may be a confounder and/or an effect modifier[232] of the effect of GBS visits on subjective well-being, but not a mediator. We hypothesized that physical activity is on the causal pathway between GBS visits and subjective well-being and is therefore a potential mediator of this relationship.[233] A number of other variables, including age, season and urban-rural status, were considered as potential confounders.[234]

Physical activity (walking)

We particularly focused on walking using the derived variable included with NSW data 'minutes walked per week qualifying as moderate activity'. Moderate intensity activity is defined as that performed at 3.0–5.9 times the intensity of rest,[235] and examples include brisk walking.[236] The NSW variable is

derived from a series of questions asking participants to report their walking frequency, duration and intensity.[228] Physical activity data were only available for our analyses for 2016/17, because in 2018/19, the NSW subsample of survey respondents who answered physical activity questions was different to the subsample answering GBS visit questions. This value was also capped at 420 minutes per week (i.e. values greater than 420 were re-coded as 420), to maintain consistency with the time outdoors variable (9% of values were recoded).

Potential confounders and effect modifiers

Covariates with available data are listed in *Table 20*. Potential confounders and effect modifiers were identified based on existing literature as described above.

The material deprivation measure is a derived binary indicator included in the NSW dataset; the measure is the same as that used in the UK Family Resources Survey.[237] It is based on an additive score summarising whether or not the participant could afford a series of items, such as 'a holiday away from home for at least a week a year'. The derived binary variable provided with NSW data classifies pensionable age adults as in material deprivation if they scored 20+ and other adults if they scored 25+ (different items used for the two age groups).

Data linkage

Individual (CMD) and household (GBS metrics, WIMD) level data were anonymously linked to NSW datasets within SAIL using ALFs (individual level) and RALFs (household level). Linkage was carried out through standard anonymised processes in SAIL (see *Chapter 2*), and permission to access selected variables from the linked survey data was obtained from the Welsh Government. The final

TABLE 20 Covariates included in modelling

Variable (role)	Data source	Description	Continuous/ categorical
Age (confounder)	NSW	Individual age at time of survey: 16–24, 25–44, 45–64, 65–79, 80+	Categorical
Economic status (confounder)	NSW	Employed; economically inactive; Unemployed (n.b. 'economically inactive' includes full-time students and retired)	Categorical (3 levels)
Gender (confounder)	NSW	Male/female (other responses too few for inclusion)	Categorical (2 levels)
Material deprivation (confounder/modifier)	NSW	In material deprivation/not in material deprivation (see text)	Categorical (2 levels)
Physical activity (walking) (mediator)	NSW (2016/17 only)	Minutes walking per week meeting moderate activity threshold	Continuous (capped at 420 mins)
Rural–urban category (confounder/modifier)	NSW	LSOA-level; various versions in NSW; version used is 3 levels: urban (> 10,000); town and fringe; village, hamlet and isolated dwellings	Categorical (3 levels)
Season (confounder)	NSW	Spring (Mar/Apr/May); summer (Jun/Jul/Aug); autumn (Sep/Oct/Nov); winter (Dec/Jan/Feb)	Categorical (4 levels)
Use of car (confounder)	NSW	Yes/No	Categorical (2 levels)
WIMD 2014 (confounder)	SAIL Reference	Quintiles of rankings of LSOA-level deprivation based on eight domains	Categorical (5 levels)
Local authority (confounder)	NSW	For proximity to nearest GBS only, local authority is also included to account for potential differences in GBS data	Categorical

sample size was 5791 after removal of missing data at each stage of the linkage process (*Appendix 6*, *Table 38*).

Statistical model development and analyses

Hypotheses

The research questions are reiterated here and hypotheses addressing them stated. Several of these include 'prerequisite hypotheses' (e.g. to investigate mediation effects). Statistical model development is described in more detail.

Question 1: Is greater residential GBS exposure associated with higher subjective well-being, and is this mediated by intentional GBS use for leisure?

Hypothesis: people with greater residential GBS exposure report higher subjective well-being scores, and this relationship is mediated by the amount of leisure time spent outdoors in GBS each week.

Prerequisite hypotheses

1. People with greater residential GBS exposure report higher subjective well-being:
    i. WEMWBS scores
    ii. life satisfaction scores.
2. People with greater residential GBS exposure spend more leisure time outdoors on visits to GBS per week.
3. People who spend more leisure time outdoors on visits to GBS each week report higher subjective well-being:
    i. WEMWBS scores
    ii. life satisfaction scores.

Question 2: Is greater residential GBS exposure associated with lower CMD likelihood, and is this relationship mediated by intentional GBS use for leisure?

Hypothesis: people with greater residential GBS exposure are less likely to be CMD cases, and this relationship is mediated by the amount of leisure time spent outdoors in GBS each week.

Prerequisite hypotheses:

1. People with greater residential GBS exposure have a lower likelihood of being a CMD case.
2. People with greater residential GBS exposure spend more leisure time outdoors per week.
3. People who spend more leisure time outdoors each week have lower likelihood of being a CMD case.

Question 3: Are associations between residential GBS exposure and use, and subjective well-being mediated by physical activity (walking)?

Hypothesis: people who spend more leisure time outdoors each week report higher subjective well-being, and this relationship is mediated by weekly duration of walking classed as moderate-vigorous physical activity.

Prerequisite hypotheses:

1. People who spend more leisure time outdoors each week tend to report higher subjective well-being:
    i. WEMWBS
    ii. life satisfaction scores.
2. People with greater residential GBS exposure tend to report longer durations of walking physical activity per week.

3. Individuals reporting more walking physical activity tend to report higher subjective well-being:
    i.   WEMWBS
    ii. Life satisfaction scores.

Question 4: Are associations between residential GBS exposure and use, and subjective well-being modified by individual-level deprivation?

Hypothesis: people with greater residential GBS exposure, or who spend more leisure time outdoors in GBS each week, report higher WEMWBS scores, and these associations are stronger among those classified as living in material deprivation compared to those not.

Statistical analyses and interpretation

Statistical analyses

Modelling approaches, including predictors and outcomes, are outlined in *Appendix 7*, *Table 39* and *Figures 45–58*. All analyses were undertaken in RStudio (version 1.4.1103; Posit Software, Boston, MA) with R (version 4.0.54). Generalised linear models (GLMs) were the primary tool for statistical analyses in this chapter. However, we first used generalised additive models (GAMs) to explore the shape of the relationship between the predictor and outcome to inform the GLMs. GAMs are an extension of GLMs, which are highly flexible and do not assume linearity in the relationships between explanatory variables and the mean of the response.[238] The GAM results and subsequent modelling decisions are presented in *Appendix 7* and briefly summarised below. However, the results from GLMs are more straightforward to interpret and hence GLM results are reported.

To test for effect modification, interaction terms were included in GLMs. To test for mediation, we used the straightforward causal steps approach as proposed by Baron and Kenny.[239] With this approach, evidence of mediation is indicated by a reduced effect size between predictor and outcome when the proposed mediator is included in the model. While in practice this approach is the same as investigating confounding, the theoretical proposition of mediators as being on the causal pathway means interpretation and implications are different.

Models were weighted using NSW sampling design weights,[102,228] which were converted to frequency weights using the 'rescale_weights()' function in the 'parameters' R package.[240] Models with subjective well-being measures (WEMWBS, life satisfaction) as outcome variables are specified as Gaussian, and those with binary outcome variables (CMD likelihood) are specified as binomial.

Potential confounders as described above were included in fully adjusted models. However, given the potential for over-adjustment, models were also run without inclusion of urban/rural category (urban/rural category correlates strongly with ambient greenness by definition).

GAM-based modelling decisions

The results from the GAMs informed further modelling. Details are in *Appendix 7* and modelling decisions are summarised here. GAMs indicated that the relationships between EVI and both subjective well-being measures are quadratic. Therefore, EVI was modelled with both linear and quadratic terms when predicting both WEMWBS and life satisfaction in subsequent GLMs. Complex non-linear associations were found between proximity to the nearest GBS and WEMWBS for categories of urban/rural and deprivation status. As such, categories were derived based on visually inspecting the relationships for both identified complex relationships (*Appendix 7*, *Figure 51*) and adjusting cut-off points to the nearest 50 m. These were 0–< 100, 100–< 300, 300–< 500 and > 500 m. Of note, the categories for proximity to nearest GBS correspond with the EVI buffer (300 m as described

in *Chapter 3*) and previous work in the literature[211] and 300 m is considered to correspond to an approximate walking distance of five minutes.[241]

There is the potential for a complex non-linear association between EVI and CMD, and to more easily translate into policy we chose to model EVI as categories when predicting CMD. Given high uncertainty due to relatively few respondents with EVI values above 0.6 (*Table 21*, *Appendix 7*, *Figure 50*) in Wales and the low number of people with a CMD living in locations where EVI > 0.6 (*n* = 26), we categorised EVI as < 0.2, 0.2–0.4 and > 0.4 in subsequent GLMs predicting CMD.

Proximity to nearest GBS from the home was modelled as a linear and quadratic term in GLM analyses with CMD as the outcome because the relationship was found to approximate a quadratic relationship for categories of urban status and deprivation status in the GAM (*Appendix 7*, *Figure 49*). The GAM indicated a quadratic relationship between time outdoors and WEMWBs, and linear relationships with life satisfaction and CMD status (*Appendix 7*, *Figures 55* and *56*).

Results

Following presentation of summary descriptive statistics for the sample, each hypothesis is addressed in turn with findings from statistical models as described above. Model results are presented in both tables and figures. Tables include the results from three modelling steps: unadjusted (key predictor(s) only); adjusted with covariates (excluding urban/rural category); and adjusted with covariates including urban status. Indicators of model fit are given by R^2 and Akaike information criterion, which takes into account the number of variables. We present only the results for key predictors below; for full model results see *Report Supplementary Material 2*. Figures are of predictions (to allow visualisation of non-linear relationships) of the outcome based on the fully adjusted GLMs (including urban status) and with fixed covariates.

Descriptive statistics
The total sample size of participants included in modelling was 5971. Nearly 60% (weighted) of participants lived in areas with an EVI of 0.2–0.4, with only 11% with an EVI of < 0.2 and 4% with EVI ≥ 0.6 (*Table 21*). The vast majority of people (82%) lived less than 300 m, and only 4% lived more than 500 m from the nearest green or blue space. A total of 40% of the sample reported spending more than two hours outdoors for leisure per week (*Table 21*). In terms of potential modifiers, 64% lived in urban areas and 13% were classified as living in material deprivation (*Table 21*, *Report Supplementary Material 2*, *Table 1*). Owing to the reduced sample for physical activity, the sample size for analyses of mediation by duration of walking physical activity was smaller at 3310 (*Table 21*). Of this sample, 63% spent no time walking at an intensity qualifying as moderate-vigorous physical activity.

Modelling results
Unless otherwise stated, modelling results described here are for fully adjusted models including urban/rural status.

Question 1: Is greater residential exposure to green and blue spaces associated with higher subjective well-being, and is this mediated by intentional use of green and blue spaces for leisure?
Results presented in *Figure 22* (full results in *Report Supplementary Material 2*, *Table 2*) show that, partially supporting our hypothesis, in unadjusted and adjusted models, EVI was associated with WEMWBS (EVI beta −10.2, 95% CI −17.13 to −3.17; EVI^2 beta 12.49, 95% CI 3.02 to 21.97). However, rather than the hypothesised positive linear relationship (greater exposure associated with better well-being), we found a U-shaped relationship such that WEMWBS decreased as EVI increased to around 0.4 but then increased with increasing EVI at values greater than 0.4. For example, an

TABLE 21 Sample descriptive statistics for selected variables for the modelling sample

Category	Full sample (n)		WEMWBS (n)[a]		Life satisfaction (n)[b]		People with CMD (%)[c]	Weekly leisure time outdoors (minutes)[d]		Weekly time walking (minutes)[e]		Walking sample (n)	
	Counts	Weighted %age	Mean	SD	Mean	SD		Mean	SD	Mean	SD	Counts	Weighted %age
Full sample	5971		50.87	9.44	7.76	1.86	14.49	136.64	156.04	89.99	150.80	3310	
EVI													
0–<0.2	679	10.75	50.89	9.21	7.70	1.82	16.04	140.49	148.07	96.23	103.04	343	8.56
0.2–<0.4	3494	57.87	51.14	9.18	7.77	1.83	14.90	145.75	158.22	88.02	103.24	1659	49.38
0.4–<0.6	1577	27.31	50.92	9.42	7.92	1.71	13.37	149.16	158.62	100.67	134.41	1120	35.73
0.6–<0.8	221	4.07	53.23	7.63	8.00	1.71	12.01	169.01	161.01	127.09	157.91	188	6.33
Proximity to nearest GBS (m)													
0–<100	1763	30.24	50.93	9.75	7.83	1.80	14.35	151.69	159.77	98.05	117.07	962	30.30
100–<300	3088	51.51	51.21	9.06	7.81	1.78	15.22	144.83	156.35	98.36	117.12	1721	51.08
300–<500	903	14.35	50.97	8.81	7.78	1.81	13.16	141.52	155.61	86.31	110.11	502	15.11
500–1100	217	3.89	52.51	7.84	7.93	1.76	10.73	161.12	159.33	77.45	94.06	125	3.51
Weekly leisure time outdoors (minutes)													
0	1782	25.51	48.73	10.27	7.44	2.19	18.50	0.00	0.00	41.49	84.61	1069	27.45
1–60	1256	21.24	51.11	8.49	7.94	1.69	15.55	35.77	17.90	76.39	106.26	714	22.10
61–120	700	13.42	51.03	8.53	7.73	1.61	11.05	97.85	19.19	105.58	114.71	409	13.88
121+	2233	39.82	52.74	8.70	8.02	1.56	12.50	317.26	108.16	144.35	126.12	1118	36.57
Weekly time walking (minutes)													
0			49.99	9.54	7.57	1.91	16.45	112.70	145.04	0.00	0.00	2145	63.36
1–60			51.53	8.20	8.19	1.78	9.69	90.74	128.64	39.04	17.74	165	4.98

continued

TABLE 21 Sample descriptive statistics for selected variables for the modelling sample *(continued)*

Category	Full sample (n) Counts	Full sample (n) Weighted %age	WEMWBS (n)[a] Mean	WEMWBS (n)[a] SD	Life satisfaction (n)[b] Mean	Life satisfaction (n)[b] SD	People with CMD (%)[c]	Weekly leisure time outdoors (minutes)[d] Mean	Weekly leisure time outdoors (minutes)[d] SD	Weekly time walking (minutes)[e] Mean	Weekly time walking (minutes)[e] SD	Walking sample (n) Counts	Walking sample (n) Weighted %age
61–120			52.43	6.85	8.24	1.30	13.70	129.16	128.58	102.57	18.19	148	4.12
121+			52.23	8.85	8.04	1.62	10.21	195.19	157.89	325.07	110.47	852	27.54
Urban status													
Urban	3899	63.60	50.92	9.26	7.77	1.81	15.81	137.82	153.55	93.02	113.04	2104	62.84
Town and fringe	810	14.23	51.27	9.12	7.75	1.84	12.30	164.14	164.93	89.42	122.63	512	16.47
Village, hamlet and isolated dwellings	1262	22.17	51.70	9.03	8.00	1.71	12.08	162.63	161.46	108.89	116.67	694	20.70
Material deprivation													
Not in material deprivation	5138	87.08	52.13	8.49	8.02	1.62	12.29	150.62	157.81	98.93	115.72	2842	86.15
In material deprivation	833	12.92	44.46	10.85	6.45	2.25	29.26	123.11	152.85	75.69	111.70	468	13.85

a WEMWBS (scale range 14–70; higher number indicates an increased sense of well-being.
b Life satisfaction (scale range 0–10; higher number indicates increased life satisfaction.
c Percentage with high likelihood of a CMD.
d Derived leisure time spent outdoors per week (minutes).
e Derived time spent walking at or above moderate activity threshold per week (minutes).
Full sample descriptive statistics, including all covariates are in *Report Supplementary Material 2, Table 1*. Note that variables modelled as continuous in further analyses are summarised here in categories (EVI, time outdoors and time walking).

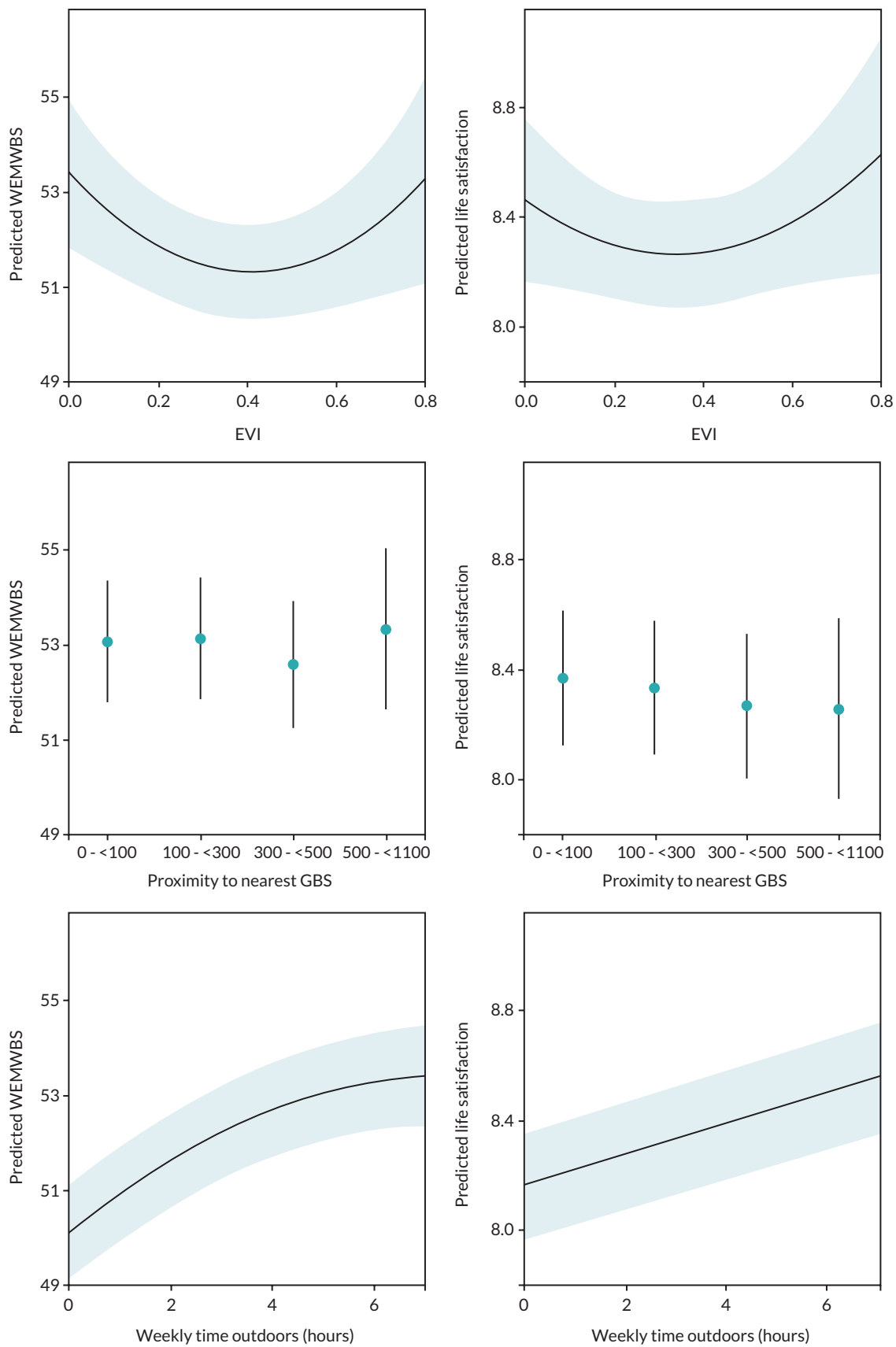**FIGURE 22** Predicted subjective well-being (WEMWBS and life satisfaction), based on fully adjusted GLMs.
Note: Predicted values are based on fixed covariates where WIMD = middle quintile, urban status = urban, age = 16–24 years, sex = female, economic status – employed, material deprivation = not in material deprivation, car use = yes, season = autumn, wave = 2016/17 and (GBS only) local authority = Cardiff.

increase in EVI from 0.2 to 0.4 was associated with an average decrease in WEMWBS of 0.53, while an increase in EVI from 0.4 to 0.6 was associated with an average increase in WEMWBS of 0.47 (not considering covariates). This pattern and coefficients were similar whether or not urban/rural category was included in models (fully adjusted without urban; EVI beta −9.69, 95% CI −16.58 to −2.79; EVI^2 beta 12.49, 95% CI 3.33 to 21.66).

A similar pattern of association was found between EVI and life satisfaction (*Figure 22*; EVI beta −1.16, 95% CI −2.51 to 0.19; EVI^2 beta 1.69, 95% CI −0.15 to 3.53). Proximity to nearest GBS was not related to either WEMWBS or life satisfaction (*Figure 22* and *Report Supplementary Material 2*, *Table 3*). There was some suggestion of decreasing WEMWBS with increasing distance to nearest GBS, but CIs were very wide. Time outdoors was positively related to WEMWBS, with evidence of diminishing marginal returns after about five hours (positive linear coefficient and negative quadratic term; *Figure 22*). For example, an increase in time spent outdoors from zero to one hour was associated with an average increase in WEMWBS of 0.83, while an increase from six to seven hours was associated with an average increase in WEMWBS of 0.12. Time outdoors was positively and linearly related to life satisfaction, although with no indication of diminishing returns (*Report Supplementary Material 2*, *Table 4*).

EVI was positively related to time outdoors, with individuals living in greener areas spending more time on visits to GBS than those living in less green areas. An increase of 0.1 EVI was associated with an average increase in weekly time outdoors of 4.34 minutes (EVI beta 43.41, 95% CI 6.85 to 79.97; *Report Supplementary Material 2*, *Table 5*). There was some suggestion that those living further away from their nearest GBS spend less time outdoors than those living closer (300–500 m vs. < 100 m beta −16.72, 95% CI −29.42 to −4.03) but no evidence of an overall association across distance categories (*Report Supplementary Material 2*, *Table 5*). This result has high uncertainty due to relatively low sample size.

Wellbeing associations with covariates

Higher well-being scores on both WEMWBS and life satisfaction were generally indicated for: males (vs. females); 65–79 years (vs. 16–24 years, with lowest scores among 45–64 years); those employed (vs. economically inactive or unemployed); individuals classified as not living in material deprivation (vs. in material deprivation); and those with access to a car (vs. no access) (*Report Supplementary Material 2*, *Tables 2–4*).

Moderation by urban/rural status

Results of models including interaction terms provided some evidence of modification of the residential GBS exposure and subjective well-being association by urban/rural status. This was particularly the case for the association between GBS proximity and life satisfaction, where, for example, living further than 500 m from the nearest GBS was associated with lower life satisfaction for town and fringe residents compared to urban residents (500–1100 m × town and fringe beta 0.66, 95% CI 1.32 to 0.00).

No other significant differences were observed by urban status between EVI and both measures of subjective well-being and between GBS proximity and WEMWBS. However, likelihood ratio tests (LRTs) comparing models with and without interaction terms indicated that models with interactions were significantly different for both GBS proximity models and for EVI predicting life satisfaction (e.g. LRT; EVI and WEMWBS, χ^2 = 11.59, *p* = 0.0.41; *Report Supplementary Material 2*, *Table 6*). In common with results of models without interaction terms, confidence intervals around coefficients are very wide, and it is difficult to discern clear patterns (*Figure 23*; full results in *Report Supplementary Material 2*, *Table 6*).
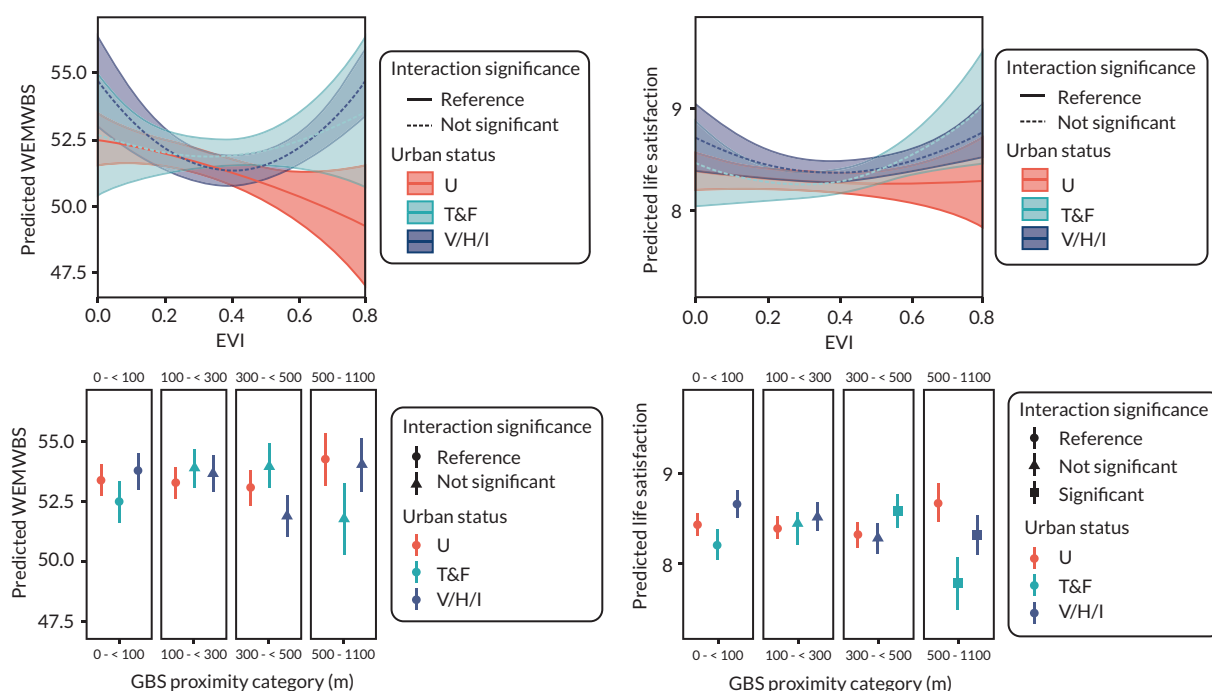
FIGURE 23 Moderation of the relationship between GBS exposures and subjective well-being by urban status. Note: T&F, town and fringe; U, urban; V/H/I, village, hamlet and isolated dwellings. Predicted values are based on fixed covariates where WIMD = middle quintile, urban status = urban, age = 16–24, sex = female, economic status – employed, material deprivation = not in material deprivation, car use = yes, season = autumn, wave = 2016/17 and (GBS only) LA = Cardiff. Significance is where $p < 0.05$ for model terms, for interaction categories this indicates that they are significantly different to the reference category.

Mediation by leisure time outdoors visiting green and blue spaces

Based on the results presented above, the prerequisite conditions for testing the mediation of GBS exposure and well-being by time outdoors are met for the relationships between EVI and both WEMWBS and life satisfaction (*Figure 24*).

To facilitate interpretation and comparison, EVI, WEMWBS, life satisfaction and leisure time outdoors were all standardised (by subtracting the mean and dividing by the standard deviation), as such the units in the results that follow are in standard deviations.

Results of the full models including terms for mediation by time outdoors are presented in *Report Supplementary Material 2*, *Tables 7* and *8*. There are non-linear relationships between both EVI and WEMWBS (U-shaped) and time outdoors and WEMWBS (diminishing returns). However, the coefficients for EVI do not change when time outdoors is added to the model (excluding time outdoors: EVI beta −0.02, 95% CI −0.05 to 0.01, EVI^2 beta 0.02, 95% CI 0.01 to 0.04; including time outdoors: EVI beta −0.03, 95% CI −0.05 to 0.00, EVI^2 beta 0.02, 95% CI 0.00 to 0.04). Therefore, these results provide no evidence of mediation of the relationship between EVI and WEMWBS by intentional use of outdoor spaces for leisure.

EVI was associated with life satisfaction with a similar effect size as WEMWBS. Similarly, the coefficient for EVI is not diminished when time outdoors is added to the model (when urban status is not included). With urban status excluded, the coefficient is reduced slightly but it is not significant with time outdoors excluded (excluding time outdoors: EVI beta −0.00, 95% CI −0.03 to 0.03; EVI^2 beta 0.02, 95% CI 0.02, 95% CI 0.00 to 0.03; including time outdoors: EVI beta 0.00, 95% CI −0.03 to 0.03, EVI^2 beta 0.01, 95% CI 0.00 to 0.03). These results therefore provide no evidence of mediation of the relationship between EVI and life satisfaction by intentional use of outdoor spaces for leisure.
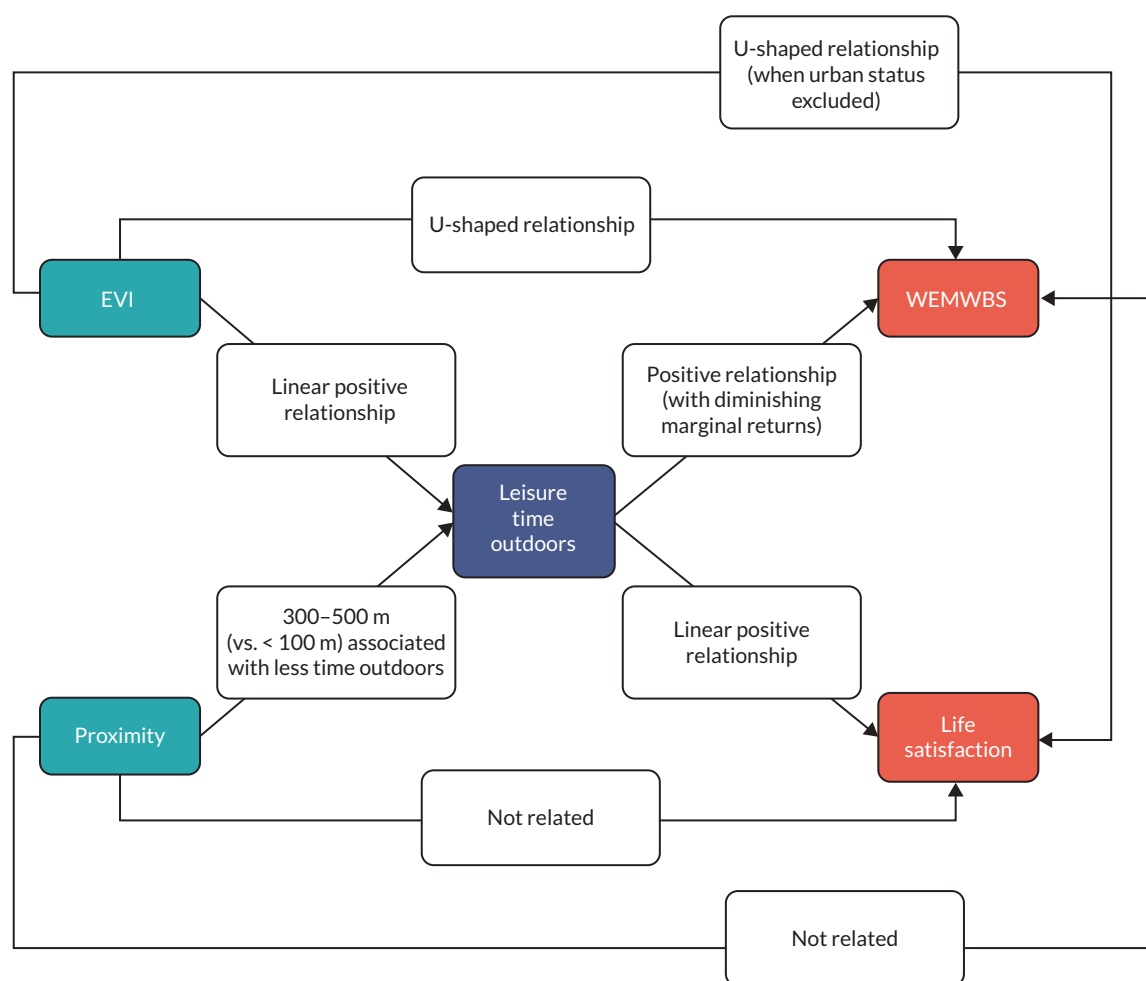
FIGURE 24 Diagram showing the prerequisite relationships for potential mediation between environmental exposures and subjective well-being outcomes when relationships are modelled separately (based on fully adjusted models; *Report Supplementary Material 2*, *Tables 2–5*).

Question 2: Is greater residential exposure to green and blue spaces associated with lower likelihood of common mental health disorders, and is this relationship mediated by intentional green and blue space use for leisure?

Neither EVI nor proximity to nearest GBS were associated with CMD (e.g. EVI 0.4 to 0.8 vs. < 0.2, OR 1.06, 95% CI 0.80 to 1.42; GBS proximity OR 1.15, 95% CI 0.97 to 1.35, GBS proximity^2 OR 0.97, 95% CI 0.95 to 1.00; *Report Supplementary Material 2*, *Table 9*), suggesting that mediation via time outdoors is unlikely. However, greater time outdoors was associated with a reduced likelihood of having a CMD, such that an increase of one hour per week is associated with a reduction in the odds of a CMD of approximately 4% (OR 0.96, 95% CI 0.93 to 0.99; *Figure 25*; *Report Supplementary Material 2*, *Table 10*). Associations between EVI, GBS proximity and CMD, and mediation by time outdoors, are depicted in *Figure 26*.

Question 3: Are associations between residential exposure to and use of green and blue spaces, and subjective well-being mediated by walking physical activity?

Residential exposure
EVI was positively associated with weekly duration of walking physical activity (PA), with a non-linear relationship indicating walking physical activity increasing from EVI values above 0.3 (EVI beta −103.28, 95% CI −260.35 to 53.79, EVI^2 beta 211.23, 95% CI 1.13 to 421.33; *Report Supplementary Material 2*, *Table 11*). For example, an increase in EVI from 0.2 to 0.3 was associated with an average
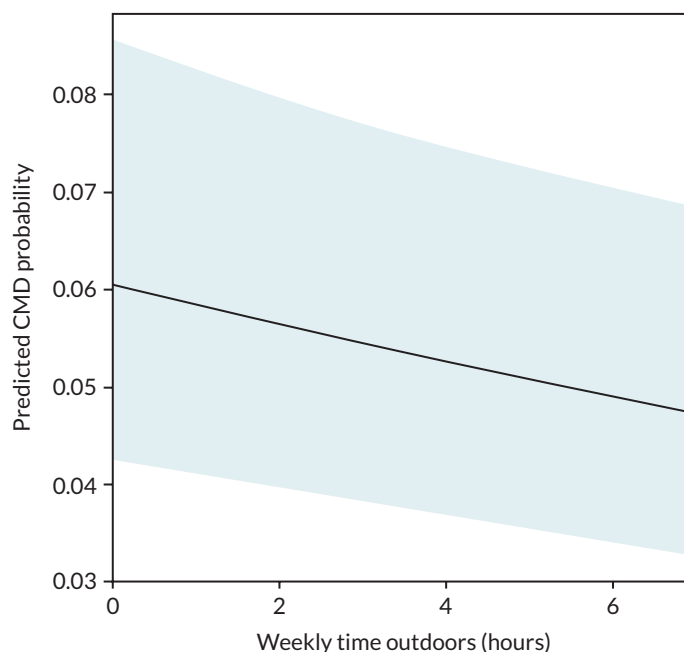
FIGURE 25 Predicted CMD by time outdoors, based on the fully adjusted model. Note: Covariates fixed where WIMD = middle quintile, urban status = urban, age = 16–24, sex = female, economic status – employed, material deprivation = not in material deprivation, car use = yes, season = autumn and wave = 2016/17.
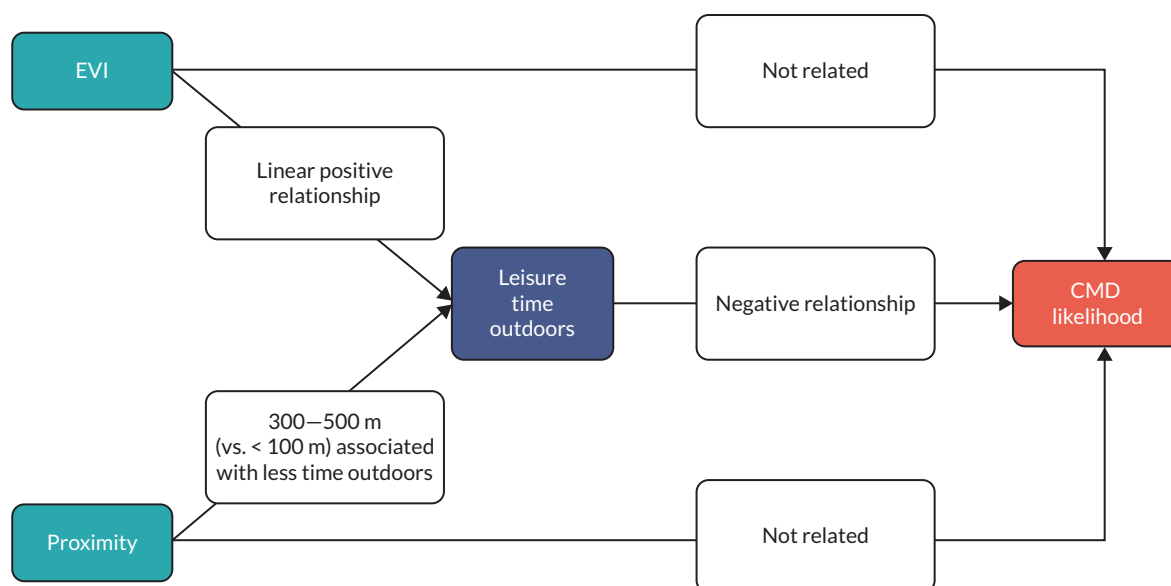
FIGURE 26 Summary of the prerequisite relationships for potential mediation between GBS exposure and CMD status by leisure time outdoors when modelled separately.

of 0.23 minutes more walking per week, while increasing from 0.3 to 0.4 was associated with 4.46 more minutes of walking per week on average. The greatest effect was found with an increase in EVI from 0.7 to 0.8, which was associated with an average of 21.36 more minutes of walking per week. Greater distance to nearest GBS was also associated with lower walking physical activity (500–1100 m vs. < 100 m beta −34.96 minutes, 95% CI −65.05 to −4.87; *Report Supplementary Material 2*, *Table 11*). Walking physical activity (hours) was positively associated with WEMWBS and life satisfaction (WEMWBS beta 0.29, 95% CI 0.17 to 0.40; life satisfaction beta 0.06, 95% CI 0.04 to 0.08; *Report Supplementary Material 2*, *Table 12*). The prerequisites for mediation as summarised in *Figure 27* (EVI related to time walking and time walking related to well-being) are therefore met for the relationship between EVI and well-being but not for proximity to nearest GBS.

FIGURE 27 Summary of the prerequisite relationships for potential mediation between GBS exposure and subjective well-being by walking as at least moderate physical activity.

For WEMWBS, the magnitude of the coefficient with EVI was not attenuated when walking physical activity was added to the model (excluding walking physical activity: standardised EVI beta −0.01, 95% CI −0.05 to 0.03, standardised EVI^2 beta 0.03, 95% CI 0.00 to 0.05; including walking physical activity: standardised EVI beta −0.01, 95% CI −0.05 to 0.03, standardised EVI^2 beta 0.03, 95% CI 0.00 to 0.05; *Report Supplementary Material 2*, *Table 13*). Similarly, for life satisfaction, EVI coefficients were not attenuated. In both cases, *p*-values for EVI coefficients increased with the addition of walking physical activity, but there was no clear evidence of mediation (i.e. negligible or no change in coefficients).

Time outdoors

We tested the hypothesis that people who spend more leisure time outdoors visiting GBS each week tend to report higher WEMWBS scores, and that this relationship is mediated by walking physical activity. Prerequisite conditions were met for potential mediation as leisure time outdoors was positively related to walking physical activity (1 minute increase in weekly time outdoors associated with 0.25 minutes greater walking physical activity on average, 95% CI 0.22 to 0.29; *Report Supplementary Material 2*, *Table 15*) and walking physical activity was associated with both WEMWBS and life satisfaction (*Report Supplementary Material 2*, *Table 12*). The time outdoors–WEMWBS coefficient was not substantially attenuated when walking physical activity was added to the model (beta reduction from 0.21 to 0.19; *Report Supplementary Material 2*, *Tables 16–17*). The time outdoors-life satisfaction coefficient was attenuated when time walking was added to the model (beta reduction from 0.07 to 0.05). As such, there is evidence of partial mediation of the relationship between time outdoors and life satisfaction by time walking but not for WEMWBS (*Figure 28*).
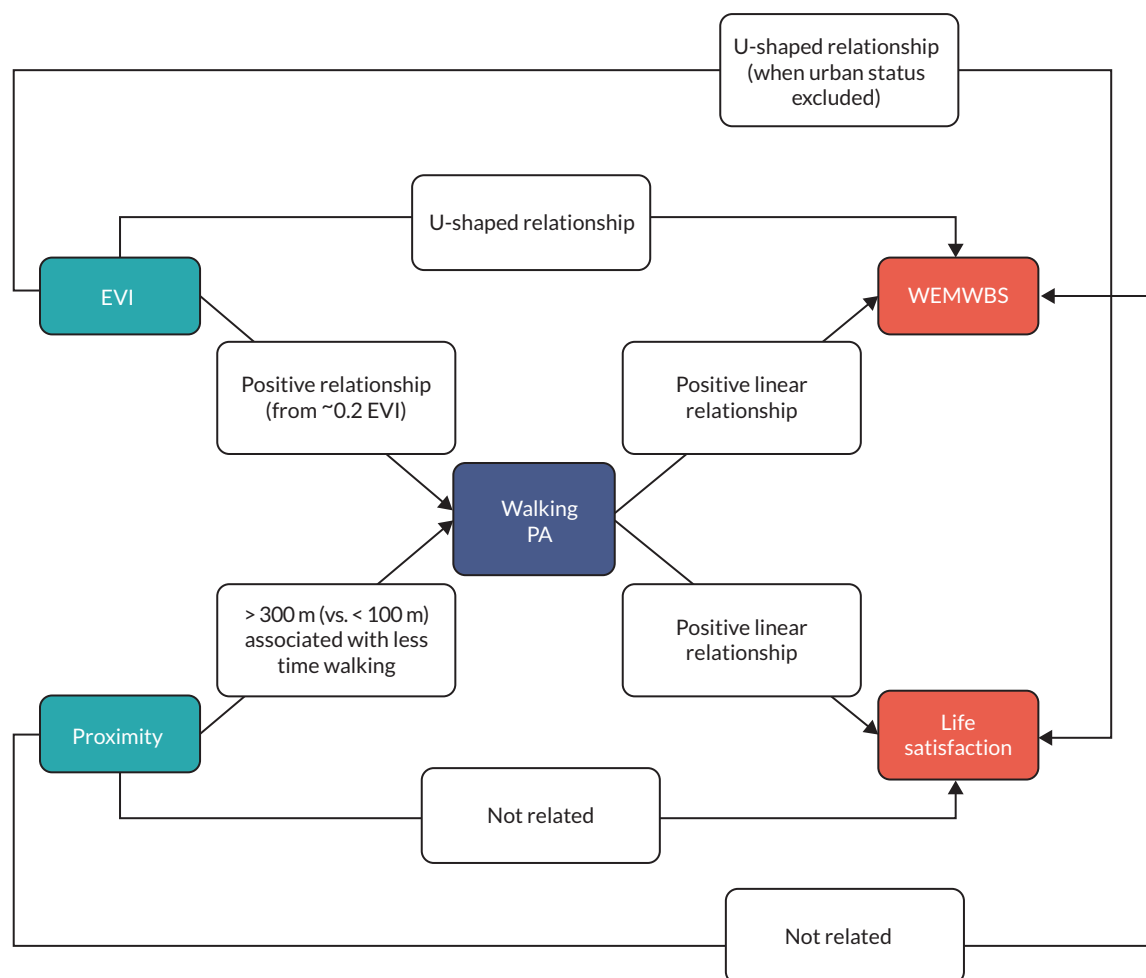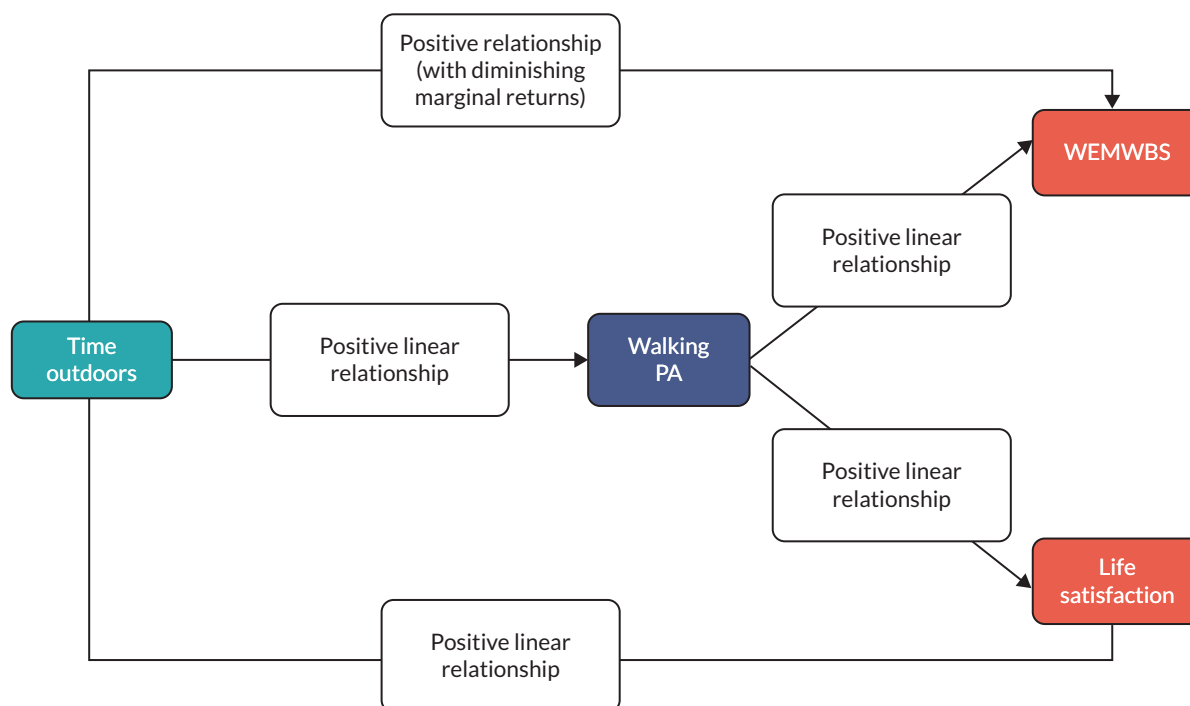
FIGURE 28 Summary of the prerequisite relationships for potential mediation between leisure time outdoors and subjective well-being by walking as at least moderate physical activity.

Question 4: Are associations between residential exposure to and use of green and blue spaces, and subjective well-being modified by individual-level deprivation?

Testing for interaction, we found no evidence of modification of the association between EVI/GBS proximity and WEMWBS or GBS proximity and life satisfaction by material deprivation (EVI predicting WEMWBS, LRT comparing model with interaction vs. without, χ^2 = 6.86, p = 0.076; *Figure 29*; *Report Supplementary Material 2*, *Table 18*). There were no significant differences by material deprivation for associations between EVI and life satisfaction, although likelihood ratios tests indicated that a model with interactions was improved (EVI predicting life satisfaction, LRT comparing model with interaction vs. without, χ^2 = 10.60, p = 0.014).

We did, however, find consistent evidence of a difference in the relationship between leisure time outdoors and both WEMWBS and life satisfaction by material deprivation (*Figure 29*; *Report Supplementary Material 2*, *Table 18*). For those in material deprivation, the association is stronger (more time outdoors is associated with higher WEMWBS) compared with those not in material deprivation up to approximately four hours/week. As *Figure 29* indicates, the deprivation gap in well-being decreases with increasing time outdoors up to that point. For people spending zero time outdoors, WEMWBS scores are, on average, 7.15 points lower among those in material deprivation compared to those not (calculations exclude covariates); at four hours per week, this gap is 4.45. Above four hours per week, the time outdoors–WEMWBS association for those in material deprivation then decreases; further increases in time outdoors are associated with lower WEMWBS relative to four hours per week. At seven hours spent outdoors per week, the difference in WEMWBS between those in material deprivation compared with those not similarly deprived is greater again at an average of 6.30 WEMWBS points. For those in material deprivation, an increase from zero to one hour of time spent outdoors is associated with an average increase in WEMWBS of 1.89, an increase from four to five hours is associated with 0.13 points increase in WEMWBS on average and an increase from six to seven hours is associated with a 0.74-point reduction in WEMWBS on average. For those not in material deprivation, WEMWBS scores increase steadily with increasing time spent outdoors but, at durations less than four hours per week, the slope is shallower than that for those in material deprivation (*Figure 29*).
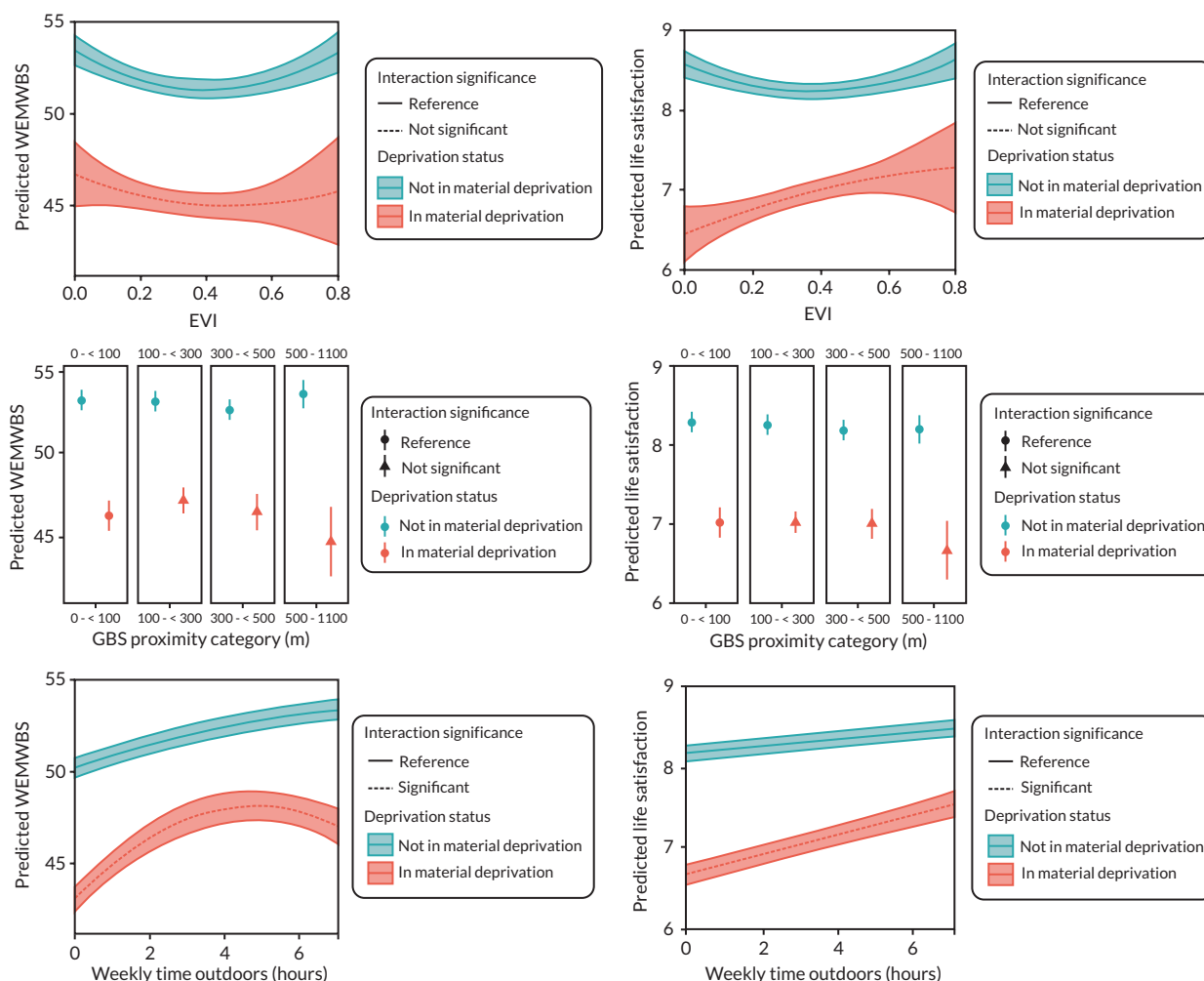
FIGURE 29 Moderation of the relationship between GBS exposures and subjective well-being by deprivation status. Predicted values are based on fixed covariates where WIMD = middle quintile, urban status = urban, age = 16–24, gender = female, economic status – employed, material deprivation = not in material deprivation, car use = yes, season = autumn, wave = 2016/17 and (GBS only) local authority = Cardiff. Significance is where $p < 0.05$ for model terms, for interaction categories this indicates that they are significantly different to the reference category.

The association between time outdoors and life satisfaction is linear, but again is stronger among those in material deprivation; that is, there is disproportionately greater benefit of time outdoors for those in material deprivation versus those not deprived. The gap in life satisfaction score between deprived and not deprived at zero hours per week is 1.53 points, whereas the deprivation gap at seven or more hours per week is 0.96 points.

Discussion

Summary of results

Question 1: Is greater residential exposure to green and blue spaces associated with higher subjective well-being, and is this mediated by intentional green and blue space use for leisure?

Ambient greenness (EVI) was associated with subjective well-being, although the association was U-shaped. Proximity to GBS was not associated with well-being. Time outdoors (intentional GBS use for leisure) was associated with WEMWBS, with diminishing returns above around four hours per week. Those spending four hours per week outdoors scored, on average, 2.6 points higher on WEMWBS

compared with those reporting no time outdoors; for those spending seven or more hours per week versus no time, the difference was 3.3 WEMWBS points (scale range 14–70). There was a linear association with life satisfaction, with four hours per week compared with no time outdoors associated with a life satisfaction score 0.2 points higher, and seven or more hours per week 0.4 higher (scale range 0–10). However, we found no evidence that time outdoors mediated the association between EVI and well-being.

Question 2: Is greater residential exposure to green and blue spaces associated with lower likelihood of common mental health disorders, and is this relationship mediated by intentional green and blue space use for leisure?

There was no association between residential GBS exposure and CMD. However, there was an inverse association between time outdoors (intentional GBS use for leisure) and CMD, with each additional hour of time outdoors associated with a 4% reduction in odds of CMD.

Question 3: Are associations between residential exposure to and use of green and blue spaces, and subjective well-being mediated by walking physical activity?

Both ambient greenness (EVI) and weekly duration of walking physical activity were associated with subjective well-being, and both ambient greenness and GBS proximity were associated with greater walking physical activity. For example, those living 500–1100 m from their nearest GBS reported, on average, around 35 minutes per week less walking physical activity than those living less than 100 m from their nearest GBS. One hour of additional weekly time outdoors was associated with around 15 minutes additional walking at or above the moderate physical activity threshold. However, we found no evidence of mediation of the EVI and well-being association by walking physical activity. We found marginal evidence of partial mediation by walking physical activity of the relationship between time outdoors and well-being. For example, one hour of time outdoors was associated with 0.05 higher life satisfaction score, but after inclusion of walking physical activity in the model, this decreased to 0.04 (scale range 0–10).

Question 4: Are associations between residential green and blue space exposure and use, and subjective well-being modified by individual-level deprivation?

We did not find evidence of modification of residential GBS exposure and well-being associations by material deprivation. However, we did find evidence for differences in the relationship between time outdoors and well-being by deprivation status. There was greater relative benefit of more time outdoors for those in material deprivation, although for WEMWBS this disproportionate benefit was primarily apparent up to four hours per week time outdoors. Among those in material deprivation, people spending four hours per week outdoors scored, on average, 4.9 points higher on WEMWBS compared with those reporting no time outdoors; for those not in material deprivation this difference was 2.2 points. For life satisfaction, associations were linear; people reporting seven or more hours per week versus no time outdoors had life satisfaction score around 0.9 points higher among those in material deprivation, and 0.3 among those not in material deprivation.

Findings in context

Residential exposure to green and blue spaces, time outdoors on visits to green and blue spaces and subjective well-being

We found a U-shaped relationship between EVI and well-being, with declining well-being as EVI increased to a value of approximately 0.4, above which well-being began to increase. We found no association between GBS proximity and well-being. However, we did find positive associations between time spent outdoors on visits to green/blue spaces and well-being, although no evidence that this mediates the EVI association.

One possible explanation for the observed U-shaped association between EVI and SWB is that low values of EVI (< 0) are produced by surface water. As indicated in *Chapter 3*, even though this analysis focuses on land area, coastal areas of Wales frequently have lower EVI values, and this is also likely to be the case for inland water bodies. Given existing evidence indicating positive associations between blue space and well-being,[38] it may be that the higher well-being values observed for individuals living in areas with very low EVI values could be explained by the positive effects of blue spaces. This effect may be non-trivial given the concentration of Wales' population in more coastal areas.

These findings are to some extent consistent with previous evidence, which is quite mixed. Some studies have found residential green/blue space to be associated with measures of mental well-being. For example, a large national-scale Canadian cross-sectional study found greenness (NDVI) based on a representative postcode was related to better mental health and well-being.[242] A 2018 systematic review suggested that there was sufficient evidence across the studies included to support a positive association between local-area greenspace and hedonic well-being (life satisfaction), but not eudaimonic well-being (life being 'worthwhile').[82] The same systematic review concluded that the evidence on green-space visit frequency was of mixed quality with inconsistent results. Of only four 'good' quality studies (all cross-sectional) investigating visit frequency included in the review, one general population study found no association with WEMWBS,[231] another also found no association, but was focused on workplace settings,[24] one general population study found a positive association with life satisfaction,[156] and another with mental health and vitality domains of SF-36.[30] The study of residents of England[156] using comparable data to the NSW survey (although using an annual visit frequency metric) found no association between residential green space and any well-being measure, but did find evidence for positive associations with visit frequency to green/blue spaces and well-being. A more recent study of four European cities using the same data as van den Berg *et al.*,[30] did not find an association between NDVI and well-being but did again find the positive association with visits to green spaces.[243] The current study is therefore consistent in findings with the latter two studies, which it most closely resembles in approach. In common with many systematic reviews relating to health impacts of green/blue space, the reviews also highlight the high degree of variability across studies, including use of a wide range of health metrics, different types of environmental data and heterogeneous scale of exposure assessment (e.g. small areas or variable sized buffer zones around home locations), making direct comparison with the existing literature difficult.[244]

A subsequent study[63] using the same England survey data found diminishing marginal returns of increased time spent visiting green/blue space. We found the association between time outdoors and WEMWBS levelled off beyond around three to four hours per week, while the England study suggested this occurs above around two hours per week.[63] However, the two analyses are based on different questions asked in the two surveys to establish time outdoors, and this is therefore not a direct comparison.

The complex association between EVI and well-being, and the lack of any evidence for mediation by visits is difficult to interpret. It is possible that other mediation pathways are in operation: neither air pollution or noise pollution were included in this study, nor was social connectedness, and these may be responsible for at least part of the relationship observed between ambient greenness and well-being.[31,245] Additionally, we could not account for the quality (either ecological or social) or specific characteristics of the GBS in this analysis.[5] Our finding of no association between proximity to nearest GBS and well-being outcomes may reflect the high proportion of people within our sample (> 80%) who lived within 300 m of their nearest GBS. Only 4% of our sample lived further than 500 m from their nearest GBS.

Residential exposure to green and blue spaces, time outdoors on visits to green and blue spaces and common mental health disorders

We found no association between residential GBS measures and CMD likelihood but did find an inverse association between time outdoors on visits to GBS and CMD likelihood. The strength of association was relatively weak, with around a 4% reduction in odds of CMD with each hour per week of time outdoors (capped at seven hours per week).

Mediation

We found no evidence of mediation of any residential GBS effects by time outdoors. However, we did find, for example, independent associations between EVI and time outdoors, and time outdoors and well-being. It is possible that the mediation analysis approach adopted here is insufficiently sensitive to detect mediation within these complex relationships. We found some marginal evidence of mediation of the time outdoors-well-being association by walking physical activity. Physical activity is proposed as one of the key mechanisms by which time spent in GBS may give rise to mental health benefits.[31] Our findings are consistent with some previous work that has suggested this process may be one of the ways in which mental health benefits accrue.[245] These mediation pathways are potentially important, since they may provide opportunities for intervention at different points along the causal pathway. In this case, this might include supporting/promoting visits to nearby GBS and 'moderately active' walking during those visits.

Modification by socioeconomic deprivation

Previous work indicates the potential for an 'equigenic' effect of exposure to green/blue space; that is, narrowing of socioeconomic health inequalities among populations exposed to more GBS.[72] A recent systematic review concluded that studies, especially in Europe, suggested that health benefits of green space were greater among people of lower socioeconomic status than those in higher socioeconomic status groups.[172]

We found no evidence of moderation of EVI-well-being associations by deprivation status, in contrast to previous studies that have found moderation of residential GBS exposure and population health associations.[3,174] However, we did find evidence for moderation of the association between time outdoors and well-being by individual-level material deprivation. This novel finding indicates that the well-being benefit of spending more time outdoors on visits to GBS accrues more rapidly for those living in material deprivation versus those not living in such circumstances. This disproportionate benefit appears to decline above around four hours per week for WEMWBS but is linear for life satisfaction. The mechanisms for this apparent disproportionate benefit cannot be determined from this analysis. However, earlier work has hypothesised that the equigenic effect could arise if positive effects of nature exposure are stronger for those living with greater stress, and weaker for those who are less in need of stress mitigation.[72] Given that stress recovery is a key proposed mechanism for psychological benefits of nature exposure,[31] it is plausible that those living with higher levels of chronic stress – such as people living in material deprivation – may have most to gain from exposures that mitigate that stress.

The deprivation gap in WEMWBS narrows by almost half for those spending four hours per week outdoors on visits to GBS versus compared with those spending no time. Interestingly, WEMWBS scores continue to climb steadily for those in the non-deprived group to the capped maximum of seven hours per week. However, for those in the deprived group, WEMWBS scores plateau and decline above four hours per week.

We can only speculate on reasons for this, and these complex findings were specifically presented to and discussed with the Health and Environment Public Engagement (HEPE) group to obtain a wider perspective for interpretation. Based on discussions among the project team and inputs from HEPE members, it is possible that many of those reporting longer durations are dog owners and walking their dogs could be (for some) more of a chore than a choice, and this may be different between deprived and non-deprived groups. The difference between deprivation groups at these higher levels of time outdoors could potentially also be associated with differences in time pressures associated with occupations, caring responsibilities and so on, making 'obligatory' time outdoors less beneficial for those whose lives are under greater time pressure. Additionally, although they are spending time outdoors, those in material deprivation may not have access to quality environments close to home, necessitating a drive if they have access to a car or dog walks through less attractive urban environments without gaining possibly greater well-being benefit of quality GBS.

The differential association with life satisfaction is simpler, with the deprivation gap (between those in material deprivation and those not) in well-being around 50% greater among those spending no time outdoors visiting GBS versus those reporting seven or more hours per week. This narrowing of the deprivation gap is again consistent with the hypothesis of disproportionate benefit, although without the complex non-linearity of the WEMWEBS association.

Strengths and limitations

This study has a number of strengths and novel aspects. The survey data provide the opportunity to investigate not only what environmental conditions are present in people's local residential area, but their visits to green and blue spaces, a key mechanism by which health benefits may actually be obtained, while the NSW sample providing data on visits to GBS is considerably smaller than that for previous, similar work for England (e.g. the data linkage with SAIL resources adds considerable value and strength to the analyses here that are not possible otherwise).[63] First, the linked data permit analysis of residential GBS exposure at much higher spatial resolution than is typically possible, with address-level geolocation coupled with road/footpath network proximity analysis. The GBS data are based on a large, integrated, national scale spatial dataset of potentially accessible spaces. Second, linkage permits assessment of CMD likelihood based on validated case-finding approaches using health service data. The data linkage adds substantially to the power of the cross-sectional survey data, and future work will be able to capitalise on this linkage for longitudinal follow-up.

Exploratory analyses using GAMs permitted the nature of associations to be assessed before implementation of more traditional GLMs, developing robust, but interpretable model results. Assessment of effect modification by socioeconomic status was able to make use of a composite individual/household-level indicator, an advantage over studies using small-area deprivation indicators.

The primary limitation of these analyses is that they are cross-sectional. Being based on recent survey data (some collected during the project), they cannot yet capitalise on the longitudinal power of the SAIL databank. Causal inference is limited, and we cannot rule out the role of reverse causality, selection bias or selective migration effects on observed associations. Our analyses of GBS proximity do not distinguish between different types, sizes or qualities of spaces, and therefore we would expect a fairly substantial degree of heterogeneity in what this exposure measure reflects, and opportunities for recreation afforded. The survey does not include data on dog ownership, which was found to be important in other studies.[246] By definition, leisure visits to GBS here include visits within Wales both from home and while on holiday, which are likely to be quite heterogeneous experiences (e.g. everyday dog walking vs. a walking holiday).

As discussed above, the EVI metric, intended to estimate ambient greenness exposure, is likely to have been impacted by the relatively high concentration of coastal and inland waters in and around Wales. Negative EVI values produced by surface water coupled with the potential for positive associations between blue space and well-being, mean that the U-shaped association observed between EVI and WEMWBS could actually indicate better well-being in 'bluer' and 'greener' areas at either end of the scale.

The sample of survey participants available for analysis was smaller than the originally planned sample specified in the protocol (c. 12,000), determined by changes to the survey data collection that evolved during the project. This impacted on the uncertainty around model estimates, with very wide confidence intervals for some key results.

In terms of specific elements of the analyses, testing for mediation was limited in several ways. Testing for mediation of residential GBS exposure associations with well-being and CMDs by intentional leisure visits, assumes that these visits are made to GBS in the immediate local area – in the case of EVI, effectively visits within 300 m. These data do not permit this assumption to be tested, but we do know from other data that a reasonably large proportion of visits are made to GBS further away from home.

For example, the Monitor of Engagement with the Natural Environment for England 2018–19[247] showed that 32% of visits were to places more than 2 miles (3.2 km) away from home, and the mean distance travelled was 5.1 miles (8.2 km).

Mediation by physical activity was restricted to considering self-reported moderate-vigorous walking physical activity as the most appropriate option available. While most visits to GBS involve walking, some do not. Also, the questions used to derive walking physical activity included walking for reasons other than leisure and in other locations, such as active travel and walking in city streets. This mediation variable is therefore an imperfect representation of what we were intending to capture (i.e. physical activity during visits to GBS).

Mediation analyses carried out here adopted a relatively simplistic, sequential modelling approach.[239] Given the complexities of the associations under investigation and potential involvement of multiple mechanisms, more sophisticated approaches such as structural equation models, or the use of g-computational procedures to deal with complex mediational relationships hypothesized[248] would be warranted to study mediation more thoroughly.[64]

Mental health status itself may act as a modifier in relationships between GBS and well-being – for example, those with depression and/or anxiety may be less likely to make visits to GBS. This is not investigated here, primarily since the relatively small sample size of the survey subsample results in a small absolute number of CMD cases. Further, this interaction/subgroup analysis was not specified a priori in the study protocol. This potential interaction is not well understood, although evidence from a 2020 18-country cross-sectional study indicates that people with depression are as likely to visit, and those with anxiety more likely to visit, than those without.[212] This evidence also indicates that our finding of lower likelihood of CMD with increasing time outdoors is unlikely to be due to reverse causality (i.e. that people with CMDs visit GBS less often).

Conclusions

These findings add considerably to the growing quantity of more robust evidence from studies investigating relationships between residential GBS, visits to GBS, and mental health and well-being. Particularly important are the findings regarding beneficial associations between time spent outdoors on leisure visits to GBS, subjective well-being and CMDs, and also the finding of a disproportionate benefit of time outdoors for the well-being of individuals living in material deprivation. The findings here do not provide simple, positive affirmation of the hypothesised relationships between residential green/blue space, mental health and well-being but are consistent with the mixed evidence base. The individual-level linked survey-and routine health data used in this study present a considerable opportunity for future environment and health research, especially if treated as baseline data for which routine data-based follow-up will permit prospective longitudinal analysis, and more sophisticated studies of mediation and moderation.

Chapter 8 Discussion

Introduction

Funding to keep parks and other outdoor natural spaces maintained or adequately provisioned as part of new housing developments is often overlooked. This is particularly true in a recession when cuts to cultural and environmental budgets in local authorities are often hardest.[249–251] The importance of access to quality GBS for maintaining mental health and well-being has been evidenced through a large volume of cross-sectional, and a growing number of longitudinal studies. However, evidence is somewhat mixed. The most robust evidence is generated using environmental and health data over a sufficiently long duration, ideally in longitudinal studies where outcomes clearly follow exposure to GBS. Our work exploited very large, data-linkage-based longitudinal data on residents in Wales for over a decade. We used two complementary quantitative approaches to investigate the impacts of GBS, and crucially changes in GBS, on mental health and well-being for our study population. Changes in exposure to local GBS arose either through moving home or from environmental changes around the home. Linkage of individual survey responses, including self-reported leisure visits to GBS, to these routinely collected health and environmental data enabled in-depth work to augment these analyses.

In this chapter, we summarise the principal findings of the study. We then discuss how the findings across our three quantitative studies integrate, using two broad headings: (1) mental health; and (2) well-being. We describe our engagement with stakeholders, before summarising the strengths and limitations of this study. Finally, we set these results in the wider context and discuss the implications for future research.

We devised an analysis plan for our funding application and developed this into a published study protocol.[83] Before data were extracted, we also wrote a detailed statistical analysis plan. However, as is common with routine data studies, the distributions of the exposure and outcome data were unknown until we extracted these and compiled descriptive statistics. After exploring the distribution of exposures and outcomes, we added a single cumulative time-aggregated exposure analysis to our project plan to fully use the spatial differences in exposures between people.

Principal findings

Our analyses have shown that:

1. Using a single time-aggregated exposure measure based on each person's residential location, people with more ambient greenness around their home, and greater access to GBS in terms of proximity, had a reduced likelihood of seeking care for any CMD during their subsequent time in the study (*Chapter 5*).
2. People who spent more time in GBS for leisure were more likely to report better subjective well-being, and the well-being benefits of spending more time outdoors for people living in material deprivation were disproportionately positive compared to people not living in material deprivation (*Chapter 7*).
3. Using multiple measures of exposure for the same individual through time, we found no evidence that increases in ambient greenness and access to GBS affected subsequent help-seeking for CMD (*Chapter 6*).

Our results add to growing evidence that exposure to GBS can be positively associated with mental health. Reviews have reported little evidence for a causal relationship between the 'greenness or blueness' of environments and adult mental health; a key limitation is lack of longitudinal studies that can more robustly examine these relationships.[5,213] At first glance, there appear to be inconsistencies

between the findings of our time-aggregated (*Chapter 5*) and longitudinal (*Chapter 6*) exposure analyses. Once we consider the structure of the data in these analyses, the differences are unsurprising. The lack of associations found in *Chapter 6* may be due to insufficient variation in greenness to detect an effect on CMD, combined with aggregating between the ability of GBS to prevent the onset of CMDs, and the role of GBS in treating CMD. Additionally, we may have captured outcomes over too short a period. For those who moved home, short-term negative impacts of a home move may be countered by immediate benefits of moving to a greener environment, rendering our results null. Over a longer timescale, positive effects may arise and be evident in the health utilisation data. Previous studies have shown that mental health improved within one year of moving to a greener area.[45]

Mental health

Our time-aggregated exposure analysis for the adult population of Wales found that people living in homes surrounded with 0.1-unit greater ambient greenness had an associated 20% reduction in the likelihood of seeking care for CMDs. People living in homes with 10% more potentially accessible spaces had an associated reduction in likelihood of seeking care of 7%. Those who lived in urban areas benefited most from increased access to GBS, potentially indicating the more acute the deprivation of 'greenness/GBS', the greater the importance to mental health. Associations between ambient greenness and potential GBS access and seeking help for CMD were not found in cross-sectional analysis of the much smaller subsample who were also in the NSW. However, when we used the linked NSW cohort to investigate more detailed behavioural data, we found that an additional hour of outdoor leisure time per week was associated with a 4% reduction in likelihood of seeking help for CMD.

The NSW analysis found a protective effect on CMD of time outdoors, consistent with previous research, which has typically investigated the relationship between GBS visits and self-reported measures of symptoms of depression and anxiety.[30,252] For the NSW subgroup we investigated the relationship using a robust, validated measure of CMD likelihood based on recorded engagement with health services. While this approach undoubtedly misses less severe cases and those that do not result in health service use, it is useful in suggesting that the benefits of spending more time visiting GBS may well be sufficient to prevent a small proportion of CMDs that require health service intervention, although this was not replicated in the time-varying analyses using the full cohort.

Well-being

For participants who completed the NSW, there was a U-shaped association between ambient greenness and subjective well-being, but no association with potential access to GBS. Time spent outdoors in GBS for leisure was associated with better subjective well-being and life satisfaction but did not appear to mediate the association with ambient greenness. People living in material deprivation showed disproportionately positive well-being benefits of spending more time outdoors. For example, among people living in material deprivation, those spending four hours per week on leisure visits to GBS scored 4.9 points higher on the WEMWBS compared with those spending no time on such visits. The difference in score for those not in material deprivation was 2.2 points. However, people in material deprivation appeared to receive diminishing benefits the longer they spent in the outdoors, while the same was not true for those not in material deprivation. This difference could be due to a range of factors, such as lower GBS quality in more deprived areas.

Economic impact

We investigated the total use of general practice services rather than restricting to seeking help for CMDs. We found no evidence that increases in GBS exposure over time and subsequent help seeking in general practice varied by socioeconomic disadvantage. Increases in greenness around the home affected seeking help for CMD in general practice for those who moved home more than once (with very slightly lower odds of seeking help for CMD). This result was not replicated for people who only moved home once, or who did not move. Therefore, we were not able to calculate any potential cost savings as the result of increased ambient greenness or potential GBS access around the home for people using general practice services.

Implications of the principal findings

Our time-aggregated exposure analyses are consistent with hypotheses of mental health benefits of GBS in the home neighbourhood (*Chapter 5*). The inconclusive results for the time-varying exposures and outcomes (*Chapter 6*) may be the result of aggregating between the ability of GBS to prevent the onset of CMDs, and the role in treating CMD.

The findings regarding time spent on leisure visits to GBS (*Chapter 7*) add considerably to a small but growing collection of studies investigating the potential mechanisms that might influence the use of GBSs, and their association with mental health and well-being. Subject to the limitations of cross-sectional analysis, our data support the positive association between time spent outdoors on leisure visits to GBS and mental health and well-being, and critically provide further evidence of a potential 'equigenic' effect of GBS on (mental) health inequalities.[72]

Our findings indicate the value of large-scale data-linkage based analyses to investigate the potential health benefits of GBS. However, they also demonstrate the significant challenges posed by the complexity of compiling time-series environmental data, which are needed to augment longitudinal population health data. Further work is needed to develop consistent, long-term, high-spatial-resolution environmental data that reliably reflect the environments that people experience over time and space. Ideally, augmented with information on combinations of GBS types that have encouraged use from diverse groups. Modelling individual behaviour in response to large scale changes in our environment may be completed prior to making large investments to help ensure health inequalities are not widened.[253,254]

We stratified our results for different age groups, however, there is likely to be residual confounding due to life stage that may not neatly correspond to age. Future analyses can investigate if the time-aggregated associations of increased GBS exposures with reduced CMD remain for subgroups based on different life circumstances. For example, we could use a 'birth in household' indicator as a proxy for becoming a new parent and compare with people without children. The data could also be organised to assess any lag effect for CMD outcomes over a longer duration. The prescription of talking therapies only may assist with stratification by case severity, enabling an assessment of the degree of effect modification.

In a previous longitudinal quasi-experimental study using the same CMD algorithm,[103] we found an association with fewer prescribed medications for those people with a history of a CMD if they lived in social housing that had been updated to national housing standards. This included intra-person differences, and a reference group living in homes that had not yet been updated (IRR 0.83, 0.70–0.97, p = 0.023).[104] This suggests that ambient greenness and potentially accessible GBS are relatively less important than the immediate home environment. The detection of associations depends on how we operationalised the environment and health data. Ambient greenness and GBS exposures may be important, but the relationships are complex and potentially differ depending on multiple factors, including the population group (old, young, minority, vulnerable, physically frail/immobile) and life circumstances (married, working, retired). It is also possible that there is not a causal relationship between exposure to GBS and CMD.

Engagement with stakeholders and the public

Using national routine data has a trade-off in excluding detailed data that might be available in local data collections. In our early stakeholder workshop (*Chapter 3*) participants highlighted their wish to include 'micro features' about the spaces. These features included the location of benches and public toilets, for example, that were unavailable in the routinely collected mapping agency data (*Chapter 3* and *Report Supplementary Material 3*). We included access points, however, that allowed us to calculate accurate

journey distances along footpaths to the nearest potentially accessible GBS. The routine mapping datasets also lack any measure of quality. Quality is a subjective measure that differs according to an individual's perspective. However, we consider that it would be useful to collate a nationally standard measure from the perspective of biodiversity, for example, to include in national mapping datasets. We found evidence that local councils collect these data but without standardisation between measures, making it impossible to create a national dataset. There may be opportunities for communities to be involved in the collection of these data, often termed as participatory GIS.[221] These have previously been scaled up to national data collections but tend to be biased towards certain participant groups.

Our patient and public involvement group helped us to interpret results from our NSW analyses, suggesting that many people reporting longer duration outdoor leisure visits could, in fact, be dog owners. For some people walking their dog could be more of a chore, exacerbated by differences in time pressures associated with occupations, caring responsibilities, and so on. They discussed that these pressures would make 'leisure' time outdoors less beneficial for those whose lives are under greater time pressure. Additionally, although they are spending time outdoors, those in material deprivation may not have access to quality environments close to home, necessitating a drive if they have access to a car or dog walks through less attractive urban environments without gaining possibly greater well-being benefit of high-quality GBS.

Challenges, strengths and limitations of the study

Our study has several strengths. The demographic dataset (WDSD), which forms the basis of our cohort, holds records for the majority (76%) of general practices in Wales.[112] The cohort is subject to minimal attrition due to the inclusion of all GP-registered individuals unless they have requested to opt out with their GP (see https://saildatabank.com/faq). By using routinely collected data, we have reduced bias (selection, participation, recall). There was also general practice health use data for most adults in Wales, giving a large number of person-years in our time series.

We included several potential confounders and effect modifiers, and our results should be generalisable to other areas in the UK and to culturally and climatically similar regions globally. We allocated exposures at a quarterly resolution and censored people who died or moved out of Wales but otherwise retained people in the study (with varying exposures). Data linkage was used to nest people in their homes and to capture their immediate environment centred on their home. We evaluated both ambient exposure and potential GBS access. In addition, explored mechanisms using a subsample of behavioural data linked at the individual level.

The individual level linked survey and routine health data used in this study presents a considerable opportunity for future environment and health research. In particular, it forms valuable baseline data for prospective longitudinal routine data collection and analysis. It also underpins sophisticated studies of mediation and moderation, and effects that vary spatially, using approaches like geographically weighted regression. Data linkage to the NWS contributed to the novelty of this study.

In brief, our study limitations were a lack of randomisation, the inability of routinely collected environmental data to give knowledge of actual exposure to GBS, time varying exposures rather than consistently allocated 'intervention and control' groups, a lack of independence between exposure measures, and limitations in the temporal sensitivity of exposure and outcome measures, or nuances in case severity to enable stratification. In addition, although we have adjusted our models for several important potential confounders, we could not repeat these investigations as thoroughly as we had intended by stratifying all analyses by these variables, due to time restrictions. Similarly, the work package 1 models were extremely resource intensive in terms of person and computing time. Experimenting with lag times would have been interesting as sensitivity analyses but would have required significant additional time and capacity. We expand on several of these strengths and limitations below, in the context of natural experimental design concepts.

A natural experiment?

We proposed a natural experiment using routine health and environment data, with a cross-sectional subcohort linked to in-depth survey data. There is an ongoing debate about the definition of a natural experiment.[255] According to the UK Medical Research Council guidance, 'By natural experiments, we mean events, interventions or policies which are not under the control of researchers, but which are amenable to research which uses the variation in exposure they generate to analyse their impact'.[256] We set out to study the impact of greening and changing GBS access on population health in a natural experimental design consistent with this perspective. Exogenous factors influence the distribution of greening and GBS access; those factors are typically external to the efforts of local communities and individuals. These would qualify our study as a natural experiment under certain definitions.[257,258] We were, however, unable to distinguish between variation in GBS due to 'events' (e.g. regeneration) or due to natural changes in vegetation or blue spaces. Moving home, a self-selection or endogenous factor, is also a component of our study. This is a larger component than we originally intended, due to the lack of viable environmental change data for GBS access. Residential mobility has been cited as solving a self-selection issue, but although we have these data, we disagree that this is a solution without understanding the reasons for selecting to move home.[259] We conclude that our study overall is observational because it does not fully meet definitions of a natural experiment. Having the right exposure/intervention data is critical for this kind of study; a different study could have developed a more typical natural experimental approach if the right data were available. However, we note that a component of our study fulfils some criteria of a natural experiment. We separated the 'movers' and 'stayers'; for the subset of people who did not move home, we have an endogenous change in naturally occurring greening that is unlikely to have been influenced by individuals or a local community but rather but wider political and economic activities. Our study is thus a longitudinal pre–post non-experimental study, with a time series of exposures, plus an embedded subcohort for whom we have behavioural data comprising actual visits to outdoor spaces.[260]

We optimised causal contrasts in our whole adult population by minimising eligibility criteria. We avoided sharp artificial boundaries that occur between sites of different exposures based on statistical areas, but instead generated individual exposures centred on their home. We strengthened our study using causal contrasts between those who moved and those who did not.[255] When designing the study, we considered adding a negative control; an additional outcome that would not be affected by the intervention. However, we were unable to identify a physical or mental health condition that could not in some way be potentially impacted by GBS. For example, the development of cancer could be influenced by the microbiome, which could be related to the presence or absence of greenness in the local environment.[261,262]

There is one randomised controlled trial where people were randomised to have adjacent vacant lots greened.[263] This rare example found that for 342 participants, depressed feelings decreased by 42% and self-reported poor mental health reduced by 63%, for those living near greened vacant lots, compared to control participants. In a commentary on the trial, Jerret and van den Bosch (2018)[259] queried whether the effect was partially due to litter clean up contributing to a tidier appearance, rather than the greening itself.[259]

For our large national study, we prepared the dataset to ensure CMD outcomes were either after the single time-aggregated exposure, or in the subsequent quarter along a series of exposures in our longitudinal analysis. Ambient greenness exposure was captured annually, and we used the data at this frequency in the longitudinal analysis, updated quarterly for people who move home, to consider outcomes for individuals who subsequently used general practice services.

In our time-aggregated exposure analysis, we examined differences between individuals in space, rather than being only restricted to changes within the same individuals through time. With careful data preparation, we ensured exposures preceded outcomes in a time-to-event/survival analysis-type approach. This provides reassurance that exposures did not come either at the same time or after

the outcome, thus reducing the potential for reverse causality. This has only previously been possible in a small number of studies. For example, a 2019 study from Denmark used NDVI at age 10 years, augmented with a cumulative NDVI measure, prior to a mental health outcome.[165] Additionally, a cohort study based in southwest England used average greenspace metrics from four follow-up waves to predict lung function at 24 years.[264] For our exposure measure we did not assume that people remained in a distinct 'allocation group'. Instead, we purposefully captured people as they transitioned between different exposures; effectively they have moved between intervention and control groups. We used a longitudinal or panel analysis, to take advantage of these transitions with the aim of assessing the impact of GBS on mental health.

We need to mention that it was not possible to establish measures of ambient greenness or access to GBS that were as robust as we would have liked; however, our analyses added significant robustness through the longitudinal approach and the time-aggregated data reduced the potential for reverse causality while maximising the spatial-temporal differences in exposure to GBS between different people at a highly granular level.

Routinely collected data

The main benefit of routinely collected health data is its historical existence at suitable granularities to enable a decade-long time series for a very large sample. The important mechanism of joining individuals to their home environment is facilitated by the presence of 'from to' dates of addresses in the demographic dataset. This allowed us to attribute individuals to home addresses for appropriate durations and therefore link exposures to each person retrospectively. Data linkage mechanisms made it possible to link health and environment data at the same spatial and temporal scales. Many previous studies on similar topics have not had coincident spatio-temporal data; many studies measuring green space exposure at a single time point, assuming that the amount of vegetation does not change over time.[265] For example, a previous longitudinal study used health data from 1991 to 2008 but had green space data from 2005 only.[4] A further study using a time series of health data from the same panel dataset, used Land Cover Map data from 2007.[4] In our study, having spatiotemporally aligned data made possible a more robust exploration of the longitudinal impact of changing environments on health.

The harmonisation of routinely collected environmental data for the whole of Wales presented challenges. We needed sufficient spatial and temporal detail to estimate exposure as precisely as possible, and to maximise the variation in GBS to give sufficient contrasting exposures through time to analyse their impact on health. We set out to generate annual or even quarterly exposures of (1) ambient greenness from satellite imagery, and (2) annual GBS access from vector data supplied by Ordnance Survey, augmented with local council map data. We scaled back our ambitions due to cloud cover over Wales, and uncertainties present in the map data from inconsistent temporal updates and changes to land use classification. We produced an annual exposure measure based on satellite imagery (using an EVI), and used only the 2018 baseline GBS access measure, capturing temporal variations in access as people moved home rather than from any in situ change in potential access. How these national data represent the reality of changing exposures, and our lack of knowledge about why people selected to move to a particular location, are important pieces of missing information that are needed to dissect these complex associations. The increase in seeking help for a CMD for people with more ambient greenness in rural areas suggests there may be a 'tipping' point where too much ambient greenness is detrimental without sufficient access to (good-quality) GBS.

To define our health outcome, we used a validated algorithm with high specificity and positive predictive value for detecting CMD (anxiety and depression) from EHRs. This validated algorithm was designed to capture the prevalence of CMD in a population. It has been successfully adapted for previous longitudinal environmental studies.[104] In the previous and current environmental health studies, we removed more severe mental health conditions, such as psychosis, before operationalising the algorithm. The algorithm uses a combination of treatments, symptoms and diagnoses as recorded by the GP to avoid counting people who use anti-anxiety or depression medications for alternative reasons. Thus, the

assignment of a CMD to a person in a particular quarter was not only based on a clinical diagnosis of a CMD. Many studies have used routinely recorded data to determine a health outcome, when obtaining a formal clinical diagnosis of a CMD would be intractable. A validation study comparing EHR algorithms (including the one we have used) to the MHI-5 demonstrated that changes in GP recording behaviour mean using only diagnosis and current treatment to identify CMD in general practice data would miss a number of true CMD cases.[103] Including historical diagnoses with current treatment, and symptoms, as in the analyses presented in this report, increases sensitivity to 0.3; we may still be missing about one-third of true cases of CMD.[103] The combination of three types of data here is a strong addition, albeit with known lack of sensitivity.

It is plausible that the severity of CMD symptoms differs between those who seek help in general practice and those who do not. The CMD measure therefore may be insufficiently sensitive to capture less severe CMD symptoms. These less severe symptoms may be more amenable to treatment or prevention from exposure to ambient greenness or access to GBS around the home than the CMDs that were captured by the algorithm. If those with milder CMDs are less likely to seek care in general practice but are also those whose CMD symptoms are more likely to respond positively to GBS, this could have diluted associations between change in GBS and milder CMDs in our study. Cognitive behavioural therapies and other non-medication treatments were not included in our CMD case definition as this information is not available in the WLGP dataset. Low sensitivity and lack of data on non-drug treatments, may have reduced the number of cases of CMD we were able to identify, with associated reductions in the precision of estimates. This may also affect the ability of EHR algorithms to capture changes in numbers of those seeking help for CMD over time, since referrals to talking therapies are increasing.[266] In future, having these data may improve the PPV for the CMD case algorithm and provide a more sensitive test of the association between GBS and CMD. Using mental health and well-being surveys linked to environmental metrics may capture CMDs where care has not (yet) been sought. This may provide greater sensitivity to detect associations between changes in the 'greenness or blueness' of the environment around the home and CMD. Cohort studies, or longitudinal panel studies, will be required to move beyond the existing evidence provided by cross-sectional studies.

The immense task of coding and data management within this project should be highlighted. It is important to ensure that each project has sufficient time to properly document the data preparation, and to ensure that code written as part of the project is made publicly accessible, as highlighted by the 2022 Goldacre and Morley review.[267]

Implications for decision-makers

Poor mental health costs the Welsh economy at least £4.8 billion annually.[268] Evidence from our research highlights the important role of adequate availability of, and access to GBSs for public health. Further, we have shown that GBSs have a role to play in reducing inequalities in mental health. We suggest our findings have a number of implications for decision makers.

Health policy and delivery

Our findings suggest that GBSs should be considered an 'asset' in efforts to promote mental health. GBSs contribute towards healthy living environments that can support population mental health, but can also be considered as a tool for health improvement strategies. For instance, as a setting for promotion of physical activity such as walking. The evidence adds to that which has justified the involvement of the natural environment sector within integrated care approaches.

Perhaps one of the most important contributions of the study to informing health policy and delivery, is the additional evidence supporting the 'equigenesis hypothesis';[269] that GBSs have a role to play in reducing socioeconomic inequalities in mental health. We showed that the well-being benefits of spending more time outdoors for people living in material deprivation were disproportionately positive

compared to people not living in material deprivation. These findings suggest that the equitable accessibility of GBSs should be integrated in future strategies to reduce socioeconomic and other inequalities in mental health. Evidence from our research supports the need to develop public health interventions to promote more leisure time spent per week visiting GBS, particularly among those in material deprivation. Care needs to be taken to ensure any strategy is equitable and recognises the interaction between the personal circumstances and attitudes of the individual, and the availability of good quality and safe accessible green or blue space resources. The evidence we provide here can be used within strategies to support the government's stated ambition: 'to narrow the gap between the richest and poorest'.[270]

Spatial planning and green infrastructure

We showed, using the single time-aggregated exposure measure based on each person's residential location, that people with more ambient greenness around their home, or greater access to GBS in terms of proximity, had a reduced likelihood of seeking care for any CMD. People living in homes with 10% more potentially accessible spaces had an associated reduction in likelihood of seeking care of 7%.

These findings support the argument that green spaces should be protected where they currently exist in neighbourhoods, and integrated, at a suitable scale, into new developments. This is particularly important in urban areas, where the association was the strongest. Green infrastructure plans and strategies should recognise the associations of greener living environments and mental health, integrating appropriate metrics of provision and accessibility. Green infrastructure decision-makers should consider the potential equigenic benefit to the mental health of those living in homes with more ambient greenness or potential access to more GBS.

Environmental management and nature conservation

Our evidence also supports the ongoing efforts to promote the use of the natural environment for benefits to mental health and other outcomes. Consistent with previous studies, the NSW analysis found a protective effect on CMD of time outdoors. We showed that each additional hour of outdoor leisure time per week was associated with a 4% reduction in likelihood of seeking help for CMD. And as noted above, people living in the greatest material deprivation benefited the most. This suggests that efforts to enable and support people to use the outdoors for leisure, including to be physically active, should continue.

Data collection and infrastructure

This study demonstrated the significant benefits of data linkage of environmental data to general practice records at a large scale. Decision makers within the health and social sector who are setting up new data infrastructures should include data linkage that is based on a flexible combination of identifiers such as name, date of birth and/or address, rather than relying solely on the assumption that each person will have an NHS number. This not only has the potential to reduce bias, but also expedites the linkage of environmental data,[271] and data from other systems outside of the health care sector, that are important to consider as part of a wider systems perspective.[272] This allows evidence to be generated to include mechanisms that are further upstream and may help to improve population health.

Data collections from surveys provide rich behaviour data for a limited number of participants. Here we demonstrated the utility of having consent to data link participant responses to their routinely collected health utilisation data. The ability to nest the 'deep, rich' data for several thousand participants within collections of 'wide, shallow' health utilisation data for the whole population is valuable. We suggest that data infrastructures should include mechanisms to allow these types of linkages. Additionally, we recommend that organisations collecting primary data request participants to consider consenting to have their responses linked to their health records. Data providers such as GPs, researchers and the public should be apprised of the value of these data linkages so they may help advocate for data infrastructures in their region to facilitate linkages, making the most of complementary collections of primary and routine data.

Access to detailed health usage data in this study was sufficient in terms of its temporal granularity. We could have studied changes in health utilisation as often as monthly if we had deemed this an appropriate temporal scale. However, we faced challenges in the availability of suitable environmental data. The temporal granularity of the environmental data was recorded insufficiently and annual changes in the vector map data could not be included in our analyses, instead we were restricted to using a single year of vector data and relying on raster data captured using satellite imagery.

Implications for future research

The results from the detailed survey data implied there is a complexity to the associations; we found different mechanisms of GBS use, and effect modification by material deprivation. We found broadly similar results for our time-aggregated analysis, but not in our time-varying analysis. We believe that using this large, powerful dataset should mean that it is possible to uncover similar associations, but that the organisation of the data in the time-varying analysis has combined too large a range of insufficiently sensitive exposures to examine these complex causal relationships. In Chapter 6, we essentially had the expectation that the exposures would have both a protective effect and be part of a treatment, without consideration of time since a house move. We suggest we could reorganise the data to include the duration since last move, and outcomes partitioned into 0–6 and 6–12 months following the move.

Future research should examine further potential effect modifiers of the association between GBS and mental health/well-being outcomes, for which there was insufficient time in this project due to significant processing time for these population-wide datasets with millions of observations. For future research it may also be beneficial to focus on a particular population subgroup hypothesised to be most sensitive to local changes in specific types of GBS. For instance, we could examine households with young children versus those without and restrict exposure to certain types of GBS, for example, access to local parks. This, combined with the availability of routine data on symptom severity, talking therapy, and in the future social prescribing, may increase the focus on people likely to be influenced by hyperlocal home environments. Additionally, it would be useful to understand how people who have a current CMD and who have potentially more accessible GBS use these spaces for recovery rather than for prevention.

This type of cohort analysis requires consistently captured, high-resolution environmental data that extends backwards (and forwards) in time. Health-focused cohort data have evolved over decades to produce extremely valuable, robust means of understanding population scale determinants of health. To complement these resources, we suggest that mapping agencies and other environmental data providers are made aware of the utility of their data for potentially health improving research, provided that their data were rendered 'backwards and forwards compatible'. That is, if improvements in the spatial accuracy of the map data, or fundamental changes to how data are captured and classified, do not unduly compromise the temporal accuracy of these data through time. This will ensure that these data capture real changes in GBS access for future longitudinal studies.

Well-being or health-related quality of life data collected routinely for patients visiting their GP would provide valuable detail to augment health service use data, having the potential to lead to targeted health interventions and economic evaluations, but are currently not collected. Thus, we used the NSW cross-sectional data that provided a wealth of detail on behaviours and well-being. These detailed data necessitate a smaller sample size. Ideally, we would have these data at multiple time points for the same individuals as part of a longitudinal cohort, although we acknowledge the cost implications of this approach. The existing survey, however, could be used to ask participants to provide information on the reasons for deciding to move home. Despite the potential for differing durations since their last move, contributing to uneven recall bias, this would allow us to calculate the degree to which there is direct confounding with selective migration associated with the exposure. By investing in urban public greenspaces, we suggest there are benefits of improved mental and physical health. There are also

additional benefits: job or food creation, biodiversity promotion, and carbon sequestration. Realising these requires a shift in the balance of decision making to place weight on protecting, enhancing and providing more appropriate greenspaces designed with local communities. Economic investments present an opportunity for sustainable transformation if funding can be leveraged to simultaneously protect and restore nature and tackle climate change and health inequalities.

Although the study period for this research preceded the current pandemic, the need for GBS was brought to the forefront during the early stages of the COVID-19 pandemic. Restrictive measures to prevent the spread of the virus, such as the instruction to exercise locally for one hour each day, unduly impacted people living in overcrowded areas of cities. There were reports of park overcrowding in some urban areas, with residents having insufficient access to GBS. Urban greenspaces may be considered a public health and social investment, providing a chance to rebalance our relationship with nature to protect against future pandemics.[273] By investing in urban public greenspaces, we suggest there are benefits of improved mental and physical health. There are also additional benefits: job or food creation, biodiversity promotion, and carbon sequestration. Realising these requires a shift in the balance of decision making to place weight on protecting, enhancing and providing more appropriate greenspaces designed with local communities. Economic investments present an opportunity for sustainable transformation if funding can be leveraged to simultaneously protect and restore nature and tackle climate change and health inequalities.

A recent National Institute for Health and Care Research (NIHR) *Public Health Research* (PHR) study led by Thompson evaluated an intervention study designed to increase access to urban woodlands.[274] The 'Woods in and around towns' study used primary data collection but, in the conclusion, the authors recommend the use of routinely collected data wherever possible. During the GBS cohort development, we would have liked a relevant intervention to embed within the routine data. This is entirely possible if data linkage systems are made available for wider public health evaluations with appropriate anonymised linkage mechanisms. Several of the investigators from both projects are part of a new UK Prevention Research Partnership, GroundsWell (https://ukprp.org/what-we-fund/groundswell). This new consortium will take a systems lens to urban GBSs, combining approaches from both these NIHR PHR projects in efforts to maximise the efficiencies from routine data, with the depth of detail collected through primary data collection. In addition, the cohort dataset is available to external researchers for collaborative research projects.[275] For further details about accessing the cohort, researchers will be able to contact saildatabank.com and Sarah Rodgers (ARCNWC@liverpool.ac.uk) for opportunities to collaborate with the original investigator team.

Conclusions

Our time-aggregated results add to growing evidence suggesting that living in areas with higher amounts of GBS has a *protective effect* on developing CMDs. These analyses have the benefit of using spatio-temporal data that are accurate to the home location and for the times people are resident, thus reducing the spatial and temporal exposure misclassifications that are common in environmental health research.

Like previous longitudinal studies, we found limited evidence for a relationship between the 'greenness and blueness' of environments and adult mental health.

Owing to the prevailing weather conditions and overall greenness of Wales, the results of this study are likely to be generalisable to other maritime temperate regions. In our survey subcohort, we tested potential mechanisms using the visit data to estimate time outdoors on visits to GBS. In support of emerging evidence, we found beneficial associations between actual time spent visiting GBS, mental health and well-being, and that these benefits may be disproportionately positive for those living in material deprivation.

Acknowledgements

Contributions of others

This study makes use of anonymised data held in the SAIL databank, which is part of the national EHR research infrastructure for Wales. It also makes use of data from the NSW. We would like to acknowledge all the data providers who make anonymised data available for research, and all the participants in the NSW. This study was advised by a steering committee chaired by Professor Steve Cummins (London School of Hygiene and Tropical Medicine) with Professor Chris Dibben (University of Edinburgh), Professor Kelvin Jordan (Keele University), Dr Paul Pilkington (University of the West of England), Ms Ruth Waters (Natural England) and Mr Gavin Atkins (MIND). We would like to acknowledge Professor Damon Berridge, whose statistical expertise was instrumental to the design of the study and who was greatly missed following his retirement due to ill health. We are grateful to Dr Kathryn Scott, Senior Grant Writer at the Centre for Drug Safety Science at the University of Liverpool who edited the report for clarity and impact. We are also grateful to the University of Exeter HEPE group and the SAIL databank consumer panel for their involvement and insights in discussions about the project and its findings in both the early and final stages.

Contributions of authors

All authors made substantial contributions to the conception or design of the work, the acquisition, analysis or interpretation of data and drafting the report or revising it critically for important intellectual content. All authors provided final approval for the version to be published.

Sarah Rodgers (https://orcid.org/0000-0002-4483-0845) (Professor of Health Informatics) led the design and execution of the study, oversaw study analyses and publications and led the preparation of the final report.

Richard Fry (https://orcid.org/0000-0002-7968-6679) (Associate Professor of GIS and Health Geographies) oversaw the preparation of the GBS exposure metrics and contributed to the design of the study and interpretation of data analyses, reporting for project management and the preparation of the final report.

Benedict Wheeler (https://orcid.org/0000-0001-9404-5936) (Senior Lecturer in Health Geography and Environmental Epidemiology) oversaw the analyses of NSW data and contributed to the design of the study and interpretation of data analyses and the final report.

Rebecca Geary (https://orcid.org/0000-0003-1417-1057) (Lecturer in Epidemiology and Public Health) contributed to the interpretation of data analyses, reporting for project management and the preparation of the final report.

Daniel Thompson (https://orcid.org/0000-0002-1769-2870) (Research Officer and Data Scientist) led the cohort preparation, data linkage for environmental and health data and final execution of the cohort for wok packages 1 and 2 (*Chapters 6*, *7* and *8*). Dan implemented the design, descriptive and exploratory data analyses for the cross sectional and longitudinal chapters, led the data linkage and analyses presentations at the Co-I meetings and study steering committees. He prepared all the tables and contributed to the writing of *Chapters 6* and *7* in the final report.

Joanne Garrett (https://orcid.org/0000-0003-0512-876X) (Research Associate in Valuing Nature Based Solutions) led the final analyses of NSW data and contributed to the final report.

Publications

Mizen A, Song J, Fry R, Akbari A, Berridge D, Parker SC, Johnson R, Lovell R, *et al.* Longitudinal access and exposure to green-blue spaces and individual-level mental health and well-being: protocol for a longitudinal, population-wide record-linked natural experiment. *BMJ Open* 2019;**9**:e027289.

Thompson DA, Geary RS, Rowney FM, Fry R, Watkins A, Wheeler BW, *et al.* Cohort profile: the green and blue spaces (GBS) and mental health in Wales e-cohort. *Int J Epidemiol* 2022;**51**(5):e285–e294. https://doi.org/10.1093/ije/dyac080

Data-sharing statement

This cohort is stored and maintained in the SAIL databank at Swansea University. This is a controlled access cohort; all proposals to use SAIL data are subject to review by an independent information governance review panel. Where access is granted, it is gained through a privacy protecting safe haven and remote access system (SAIL gateway). The cohort data will be available for collaborative research projects after 2023. For further details about accessing the cohort, contact saildatabank.com, or contact arcnwc@liverpool.ac.uk to discuss working with the original cohort developers.

References

1. Keniger LE, Gaston KJ, Irvine KN, Fuller RA. What are the benefits of interacting with nature? *Int J Environ Res Public Health* 2013;**10**:913–35. https://doi.org/10.3390/ijerph10030913

2. Nieuwenhuijsen MJ, Kruize H, Gidlow C, Andrusaityte S, Antó JM, Basagaña X, *et al.* Positive health effects of the natural outdoor environment in typical populations in different regions in Europe (PHENOTYPE): a study programme protocol. *BMJ Open* 2014;**4**:e004951. https://doi.org/10.1136/bmjopen-2014-004951

3. Wheeler BW, Lovell R, Higgins SL, White MP, Alcock I, Osborne NJ, *et al.* Beyond greenspace: an ecological study of population general health and indicators of natural environment type and quality. *Int J Health Geogr* 2015;**14**:17. https://doi.org/10.1186/s12942-015-0009-5

4. White MP, Alcock I, Wheeler BW, Depledge MH. Would you be happier living in a greener urban area? A fixed-effects analysis of panel data. *Psychol Sci* 2013;**24**:920–8. https://doi.org/10.1177/0956797612464659

5. Houlden V, Weich S, Albuquerque JP de, Jarvis S, Rees K. The relationship between greenspace and the mental wellbeing of adults: a systematic review. *PLoS One* 2018;**13**:e0203000. https://doi.org/10.1371/journal.pone.0203000

6. Kumar P, Druckman A, Gallagher J, Gatersleben B, Allison S, Eisenman TS, *et al.* The nexus between air pollution, green infrastructure and human health. *Environ Int* 2019;**133**:105181. https://doi.org/10.1016/j.envint.2019.105181

7. Van Renterghem T. Towards explaining the positive effect of vegetation on the perception of environmental noise. *Urban Forest Urban Green* 2019;**40**:133–44. https://doi.org/10.1016/j.ufug.2018.03.007

8. Gascon M, Triguero-Mas M, Martínez D, Dadvand P, Forns J, Plasència A, *et al.* Mental health benefits of long-term exposure to residential green and blue spaces: a systematic review. *Int J Environ Res Public Health* 2015;**12**:4354–79. https://doi.org/10.3390/ijerph120404354

9. Astell-Burt T, Feng X. Association of urban green space with mental health and general health among adults in Australia. *JAMA Netw Open* 2019;**2**(7):e198209. https://doi.org/10.1001/jamanetworkopen.2019.8209

10. Rodgers SE, Demmler JC, Dsilva R, Lyons RA. Protecting health data privacy while using residence-based environment and demographic data. *Health Place* 2012;**18**:209–17. https://doi.org/10.1016/j.healthplace.2011.09.006

11. World Health Organization. *10 Facts on Mental Health*. 2019. URL: https://www.who.int/news-room/facts-in-pictures/detail/mental-health (accessed 12 August 2022).

12. Vik MH, Carlquist E. Measuring subjective well-being for policy purposes: the example of well-being indicators in the WHO 'Health 2020' framework. *Scand J Public Health* 2018;**46**:279–86. https://doi.org/10.1177/1403494817724952

13. The WHOQOL Group. The World Health Organization quality of life assessment (WHOQOL): position paper from the World Health Organization. *Soc Sci Med* 1995;**41**:1403–9. https://doi.org/10.1016/0277-9536(95)00112-K

14. Diener E, Chan MY. Happy people live longer: subjective well-being contributes to health and longevity. *Appl Psychol Health Well-Being* 2011;**3**:1–43. https://doi.org/10.1111/j.1758-0854.2010.01045.x

15. Steptoe A, Deaton A, Stone AA. Subjective wellbeing, health, and ageing. *Lancet* 2015;**385**:640–8. https://doi.org/10.1016/S0140-6736(13)61489-0

16. Vos T, Barber RM, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, *et al.* Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet* 2015;**386**:743–800. https://doi.org/10.1016/S0140-6736(15)60692-4

17. Vigo D, Thornicroft G, Atun R. Estimating the true global burden of mental illness. *Lancet Psychiatry* 2016;**3**:171–8. https://doi.org/10.1016/S2215-0366(15)00505-2

18. Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, *et al.* Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet* 2013;**382**:1575–86. https://doi.org/10.1016/S0140-6736(13)61611-6

19. Stansfeld S, Clark C, Bebbington P, King M, Jenkins R, Hinchliffe S. Common mental disorders. In McManus S, Bebbington P, Jenkins R, Brugha T, editors. *Mental Health and Wellbeing in England: Adult Psychiatric Morbidity Survey 2014*. Leeds: NHS Digital; 2016.

20. Welsh Government. *Welsh Health Survey 2015*. 2016. URL: https://webarchive.nationalarchives.gov.uk/20180412105518/http://gov.wales/statistics-and-research/welsh-health-survey/?lang=en (Accessed 12 August 2022).

21. Mental Health Foundation. *Fundamental Facts About Mental Health 2016*. London: Mental Health Foundation; 2016.

22. Natural Resources Wales. *Local Green Spaces* 2020.

23. Taylor L, Hochuli DF. Defining greenspace: multiple uses across multiple disciplines. *Landsc Urban Plann* 2017;**158**:25–38. https://doi.org/10.1016/j.landurbplan.2016.09.024

24. Gilchrist K, Brown C, Montarzino A. Workplace settings and wellbeing: greenspace use and views contribute to employee wellbeing at peri-urban business sites. *Landsc Urban Plann* 2015;**138**:32–40. https://doi.org/10.1016/j.landurbplan.2015.02.004

25. Ohly H, White MP, Wheeler BW, Bethel A, Ukoumunne OC, Nikolaou V, *et al.* Attention restoration theory: a systematic review of the attention restoration potential of exposure to natural environments. *J Toxicol Environ Health Part B* 2016;**19**:305–43. https://doi.org/10.1080/10937404.2016.1196155

26. Kondo MC, Jacoby SF, South EC. Does spending time outdoors reduce stress? A review of real-time stress response to outdoor environments. *Health Place* 2018;**51**:136–50. https://doi.org/10.1016/j.healthplace.2018.03.001

27. van den Berg M, Wendel-Vos W, van Poppel M, Kemper H, van Mechelen W, Maas J. Health benefits of green spaces in the living environment: a systematic review of epidemiological studies. *Urban Forest Urban Green* 2015;**14**:806–16. https://doi.org/10.1016/j.ufug.2015.07.008

28. de Vries S, van Dillen SME, Groenewegen PP, Spreeuwenberg P. Streetscape greenery and health: stress, social cohesion and physical activity as mediators. *Soc Sci Med* 2013;**94**:26–33. https://doi.org/10.1016/j.socscimed.2013.06.030

29. Triguero-Mas M, Donaire-Gonzalez D, Seto E, Valentín A, Martínez D, Smith G, *et al.* Natural outdoor environments and mental health: stress as a possible mechanism. *Environ Res* 2017;**159**:629–38. https://doi.org/10.1016/j.envres.2017.08.048

30. van den Berg M, van Poppel M, van Kamp I, Andrusaityte S, Balseviciene B, Cirach M, *et al.* Visiting green space is associated with mental health and vitality: a cross-sectional study in four european cities. *Health Place* 2016;**38**:8–15. https://doi.org/10.1016/j.healthplace.2016.01.003

31. Markevych I, Schoierer J, Hartig T, Chudnovsky A, Hystad P, Dzhambov AM, *et al.* Exploring pathways linking greenspace to health: theoretical and methodological guidance. *Environ Res* 2017;**158**:301–17. https://doi.org/10.1016/j.envres.2017.06.028

32. Thompson Coon J, Boddy K, Stein K, Whear R, Barton J, Depledge MH. Does participating in physical activity in outdoor natural environments have a greater effect on physical and mental wellbeing than physical activity indoors? A systematic review. *Environ Sci Technol* 2011;**45**:1761–72. https://doi.org/10.1021/es102947t

33. McEachan RRC, Yang TC, Roberts H, Pickett KE, Arseneau-Powell D, Gidlow CJ, *et al.* Availability, use of, and satisfaction with green space, and children's mental wellbeing at age 4 years in a multicultural, deprived, urban area: results from the Born in Bradford cohort study. *Lancet Planet Health* 2018;**2**:e244–54. https://doi.org/10.1016/S2542-5196(18)30119-0

34. Astell-Burt T, Mitchell R, Hartig T. The association between green space and mental health varies across the lifecourse. A longitudinal study. *J Epidemiol Community Health* 2014;**68**:578–83. https://doi.org/10.1136/jech-2013-203767

35. Bolte G, Nanninga S, Dandolo L. Sex/gender differences in the association between residential green space and self-rated health: a sex/gender-focused systematic review. *Int J Environ Res Public Health* 2019;**16**:4818. https://doi.org/10.3390/ijerph16234818

36. White M, Smith A, Humphryes K, Pahl S, Snelling D, Depledge M. Blue space: the importance of water for preference, affect, and restorativeness ratings of natural and built scenes. *J Environ Psychol* 2010;**30**:482–93. https://doi.org/10.1016/j.jenvp.2010.04.004

37. Grellier J, White MP, Albin M, Bell S, Elliott LR, Gascón M, *et al.* BlueHealth: a study programme protocol for mapping and quantifying the potential benefits to public health and well-being from Europe's blue spaces. *BMJ Open* 2017;**7**:1–10. https://doi.org/10.1136/bmjopen-2017-016188

38. Gascon M, Zijlema W, Vert C, White MP, Nieuwenhuijsen MJ. Outdoor blue spaces, human health and well-being: a systematic review of quantitative studies. *Int J Hyg Environ Health* 2017;**220**:1207–21. https://doi.org/10.1016/j.ijheh.2017.08.004

39. Alcock I, White MP, Lovell R, Higgins SL, Osborne NJ, Husk K, *et al.* What accounts for 'England's green and pleasant land'? A panel data analysis of mental health and land cover types in rural England. *Landsc Urban Plann* 2015;**142**:38–46. https://doi.org/10.1016/j.landurbplan.2015.05.008

40. Nutsford D, Pearson AL, Kingham S, Reitsma F. Residential exposure to visible blue space (but not green space) associated with lower psychological distress in a capital city. *Health Place* 2016;**39**:70–8. https://doi.org/10.1016/j.healthplace.2016.03.002

41. de Vries S, ten Have M, van Dorsselaer S, van Wezep M, Hermans T, de Graaf R. Local availability of green and blue space and prevalence of common mental disorders in the Netherlands. *BJPsych Open* 2016;**2**:366–72. https://doi.org/10.1192/bjpo.bp.115.002469

42. Völker S, Kistemann T. The impact of blue space on human health and well-being – Salutogenetic health effects of inland surface waters: a review. *Int J Hyg Environ Health* 2011;**214**:449–60. https://doi.org/10.1016/j.ijheh.2011.05.001

43. Pasanen TP, White MP, Wheeler BW, Garrett JK, Elliott LR. Neighbourhood blue space, health and wellbeing: the mediating role of different types of physical activity. *Environ Int* 2019;**131**:105016. https://doi.org/10.1016/j.envint.2019.105016

44. Garrett JK, Clitherow TJ, White MP, Wheeler BW, Fleming LE. Coastal proximity and mental health among urban adults in England: the moderating effect of household income. *Health Place* 2019;**59**:102200. https://doi.org/10.1016/j.healthplace.2019.102200

45. Alcock I, White MP, Wheeler BW, Fleming LE, Depledge MH. Longitudinal effects on mental health of moving to greener and less green urban areas. *Environ Sci Technol* 2014;**48**:1247–55. https://doi.org/10.1021/es403688w

46. Lovell R, White M, Wheeler BW, Taylor T, Elliott LR. *Health and Wellbeing Evidence Review: Green Infrastructure Standards*. Defra, Natural England, PHE, MHCLG; 2019.

47. Reid CE, Clougherty JE, Shmool JLC, Kubzansky LD. Is all urban green space the same? A comparison of the health benefits of trees and grass in New York city. *Int J Environ Res Public Health* 2017;**14**:1411. https://doi.org/10.3390/ijerph14111411

48. Wyles KJ, White MP, Hattam C, Pahl S, King H, Austen M. Are some natural environments more psychologically beneficial than others? The importance of type and quality on connectedness to nature and psychological restoration. *Environ Behav* 2019;**51**:111–43. https://doi.org/10.1177/0013916517738312

49. Genter C, Roberts A, Richardson J, Sheaff M. The contribution of allotment gardening to health and wellbeing: a systematic review of the literature. *Br J Occup Therapy* 2015;**78**:593–605. https://doi.org/10.1177/0308022615599408

50. Soga M, Gaston KJ, Yamaura Y. Gardening is beneficial for health: a meta-analysis. *Prevent Med Rep* 2017;**5**:92–9. https://doi.org/10.1016/j.pmedr.2016.11.007

51. de Bell S, White M, Gri A, Darlow A, Taylor T, Wheeler B, *et al.* Spending time in the garden is positively associated with health and wellbeing: results from a national survey in England. *Landsc Urban Plann* 2020;**200**:103836. https://doi.org/10.1016/j.landurbplan.2020.103836

52. Nguyen P-Y, Astell-Burt T, Rahimi-Ardabili H, Feng X. Green space quality and health: a systematic review. *Int J Environ Res Public Health* 2021;**18**:11028. https://doi.org/10.3390/ijerph182111028

53. Zhang Y, Van den Berg AE, Van Dijk T, Weitkamp G. Quality over quantity: contribution of urban green space to neighborhood satisfaction. *Int J Environ Res Public Health* 2017;**14**:535. https://doi.org/10.3390/ijerph14050535

54. Cohen DA, Han B, Isacoff J, Shulaker B, Williamson S, Marsh T, *et al.* Impact of park renovations on park use and park-based physical activity. *J Phys Act Health* 2015;**12**:289–95. https://doi.org/10.1123/jpah.2013-0165

55. Weimann H, Rylander L, van den Bosch MA, Albin M, Skärbäck E, Grahn P, *et al.* Perception of safety is a prerequisite for the association between neighbourhood green qualities and physical activity: results from a cross-sectional study in Sweden. *Health Place* 2017;**45**:124–30. https://doi.org/10.1016/j.healthplace.2017.03.011

56. Krefis A, Augustin M, Schlünzen K, Oßenbrügge J, Augustin J. How does the urban environment affect health and well-being? A systematic review. *Urban Sci* 2018;**2**:21. https://doi.org/10.3390/urbansci2010021

57. Cox DTC, Shanahan DF, Hudson HL, Plummer KE, Siriwardena GM, Fuller RA, *et al.* Doses of neighborhood nature: the benefits for mental health of living with nature. *BioScience* 2017;**67**:147–55. https://doi.org/10.1093/biosci/biw173

58. Wood E, Harsant A, Dallimer M, Cronin de Chavez A, McEachan RRC, Hassall C. Not all green space is created equal: biodiversity predicts psychological restorative benefits from urban green space. *Front Psychol* 2018;**9**:2320. https://doi.org/10.3389/fpsyg.2018.02320

59. Lovell R, Wheeler BW, Higgins SL, Irvine KN, Depledge MH. A systematic review of the health and well-being benefits of biodiverse environments. *J Toxicol Environ Health Part B* 2014;**17**:1–20. https://doi.org/10.1080/10937404.2013.856361

60. Van den Berg AE, Jorgensen A, Wilson ER. Evaluating restoration in urban green spaces: does setting type make a difference? *Landsc Urban Plann* 2014;**127**:173–81. https://doi.org/10.1016/j.landurbplan.2014.04.012

61. Hartig T, Mitchell RJ, de Vries S, Frumkin H. Nature and health. *Annu Rev Public Health* 2014;**35**:207–28. https://doi.org/10.1146/annurev-publhealth-032013-182443

62. Coldwell DF, Evans KL. Visits to urban green-space and the countryside associate with different components of mental well-being and are better predictors than perceived or actual local urbanisation intensity. *Landsc Urban Plann* 2018;**175**:114–22. https://doi.org/10.1016/j.landurbplan.2018.02.007

63. White MP, Alcock I, Grellier J, Wheeler BW, Hartig T, Warber SL, *et al.* Spending at least 120 minutes a week in nature is associated with good health and wellbeing. *Sci Rep* 2019;**9**:1–11. https://doi.org/10.1038/s41598-019-44097-3

64. Dzhambov AM, Browning MHEM, Markevych I, Hartig T, Lercher P. Analytical approaches to testing pathways linking greenspace to health: a scoping review of the empirical literature. *Environ Res* 2020;**186**:109613. https://doi.org/10.1016/j.envres.2020.109613

65. Suppakittpaisarn P, Jiang X, Sullivan WC. Green infrastructure, green stormwater infrastructure, and human health: a review. *Curr Landsc Ecol Rep* 2017;**2**:96–110. https://doi.org/10.1007/s40823-017-0028-y

66. Dzhambov AM, Dimitrova DD. Urban green spaces' effectiveness as a psychological buffer for the negative health impact of noise pollution: a systematic review. *Noise Health* 2014;**16**:157–65. https://doi.org/10.4103/1463-1741.134916

67. Jones A, Hillsdon M, Coombes E. Greenspace access, use, and physical activity: understanding the effects of area deprivation. *Prev Med* 2009;**49**:500–5. https://doi.org/10.1016/j.ypmed.2009.10.012

68. Boyd F, White MP, Bell SL, Burt J. Who doesn't visit natural environments for recreation and why: a population representative analysis of spatial, individual and temporal factors among adults in England. *Landsc Urban Plann* 2018;**175**:102–13. https://doi.org/10.1016/j.landurbplan.2018.03.016

69. Ferguson M, Roberts HE, McEachan RRC, Dallimer M. Contrasting distributions of urban green infrastructure across social and ethno-racial groups. *Landsc Urban Plann* 2018;**175**:136–48. https://doi.org/10.1016/j.landurbplan.2018.03.020

70. Public Health Wales Observatory. *Measuring Inequalities: Overview*. 2016. URL: https://phw.nhs.wales/services-and-teams/observatory/data-and-analysis/measuring-inequalities-2016 (accessed 8 April 2020).

71. Public Health England. *Inequalities in Health*. 2018. URL: https://www.gov.uk/government/publications/health-profile-for-england-2018/chapter-5-inequalities-in-health (accessed 8 April 2020).

72. Mitchell RJ, Richardson EA, Shortt NK, Pearce JR. Neighborhood environments and socioeconomic inequalities in mental well-being. *Am J Prev Med* 2015;**49**:80–4. https://doi.org/10.1016/j.amepre.2015.01.017

73. Wheeler BW, White M, Stahl-Timmins W, Depledge MH. Does living by the coast improve health and well-being. *Health Place* 2012;**18**:1198–201. https://doi.org/10.1016/j.healthplace.2012.06.015

74. World Health Organization. *Determinants of Health*. 2020. URL: https://www.who.int/news-room/questions-and-answers/item/determinants-of-health (accessed 16 May 2023).

75. Welsh Government. *Well-being of Future Generations (Wales) Act 2015*. Cardiff: Welsh Government; 2016. URL: https://www.futuregenerations.wales/about-us/future-genera-tions-act (accessed 19 April 2023).

76. Welsh Government. *Planning policy Wales: Edition 11*. Cardiff: Welsh Government; 2018 (updated 2021). URL: https://www.gov.wales/planning-policy-wales. (accessed 19 April 2023).

77. Welsh Assembly Government. *Creating an Active Wales*. 2006. URL: http://www.wales.nhs.uk/documents/100121activewalesen.pdf (accessed 12 August 2022).

78. Natural Resources Wales. *The State of Natural Resources report 2016*. 2016. URL: https://naturalresources.wales/evidence-and-data/research-and-reports/the-state-of-natural-re-sources-report-assessment-of-the-sustainable-management-of-natural-resources/?lang=en (accessed 8 April 2020).

79. Natural Resources Wales. *Managing Today's Natural Resources for Tomorrow's Generations*. Cardiff: Natural Resources Wales; 2018. URL: https://naturalresources.wales/media/684542/final-cor-porate-plan-english.pdf (accessed 19 April 2023).

80. Natural Resources Wales. *Improving Access for All*. Cardiff: Natural Resources Wales; 2020. URL: https://naturalresources.wales/days-out/recreation-and-access-policy-advice-and-guidance/managing-access/improving-access-for-all/?lang=en (accessed 8 April 2020).

81. HM Government. *A Green Future: Our 25 Year Plan to Improve the Environment*. London: Defra; 2018.

82. Houlden V, Weich S, Jarvis S. A cross-sectional analysis of green space prevalence and mental wellbeing in England. *BMC Public Health* 2017;**17**:460. https://doi.org/10.1186/s12889-017-4401-x

83. Mizen A, Song J, Fry R, Akbari A, Berridge D, Parker SC, *et al.* Longitudinal access and exposure to green-blue spaces and individual-level mental health and well-being: protocol for a longitu-dinal, population-wide record-linked natural experiment. *BMJ Open* 2019;**9**:1–10. https://doi.org/10.1136/bmjopen-2018-027289

84. Annerstedt M, Östergren P-O, Björk J, Grahn P, Skärbäck E, Währborg P. Green qualities in the neighbourhood and mental health – results from a longitudinal cohort study in Southern Sweden. *BMC Public Health* 2012;**12**:337. https://doi.org/10.1186/1471-2458-12-337

85. Masters R, Anwar E, Collins B, Cookson R, Capewell S. Return on investment of public health interventions: a systematic review. *J Epidemiol Community Health* 2017;**71**:827–34. https://doi.org/10.1136/jech-2016-208141

86. Chen X. Urban nature's health effects and monetary valuation: a systematic review. *J Environ Plan Man* 2016;**63**:1716–37. https://doi.org/10.1080/09640568.2019.1689107.

87. World Health Organization Regional Office for Europe. *Urban Green Space Interventions and Health: A Review of Impacts and Effectiveness*. Copenhagen: WHO Regional Office for Europe; 2017. URL: https://www.who.int/europe/publications/m/item/urban-green-space-interven-tions-and-health--a-review-of-impacts-and-effectiveness.-full-report (accessed 19 April 2023).

88. Wolf KL, Measells MK, Grado SC, Robbins AST. Economic values of metro nature health bene-fits: a life course approach. *Urban For Urban Green* 2015;**14**:694–701. https://doi.org/10.1016/j.ufug.2015.06.009

89. White MP, Elliott LR, Taylor T, Wheeler BW, Spencer A, Bone A, *et al.* Recreational physical activity in natural environments and implications for health: a population based cross-sectional study in England. *Prev Med* 2016;**91**:383–8. https://doi.org/10.1016/j.ypmed.2016.08.023

90. Papathanasopoulou E, White MP, Hattam C, Lannin A, Harvey A, Spencer A. Valuing the health benefits of physical activities in the marine environment and their importance for marine spatial planning. *Marine Policy* 2016;**63**:144–52. https://doi.org/10.1016/j.marpol.2015.10.009

91. Cavill N, Gower R, Williams S. *Economic Assessment of the Health Benefits of Walking on the Wales Coast Path*. Cardiff: Natural Resources Wales; 2014.

92. Natural Resources Wales. *Evidence Needs*. Cardiff: Natural Resources Wales; 2019. URL: https://naturalresources.wales/evidence-and-data/research-and-reports/state-of-natural-re-sources-interim-report-2019/evidence-needs/?lang=en (accessed 8 April 2020).

93. Lyons RA, Jones KH, John G, Brooks CJ, Verplancke J-P, Ford DV, *et al.* The SAIL databank: linking multiple health and social care datasets. *BMC Med Inform Decis Mak* 2009;**9**:3. https://doi.org/10.1186/1472-6947-9-3

94. Ford DV, Jones KH, Verplancke J-P, Lyons RA, John G, Brown G, *et al.* The SAIL Databank: building a national architecture for e-health research and evaluation. *BMC Health Serv Res* 2009;**9**:157. https://doi.org/10.1186/1472-6963-9-157

95. *Welsh Index of Multiple Deprivation*. Welsh Government; 2022. URL: https://gov.wales/welsh-index-multiple-deprivation (accessed 8 November 2021).

96. Office for National Statistics. *2011 Rural/Urban Classification*. 2016. URL: https://www.ons.gov.uk/methodology/geography/geographicalproducts/ruralurbanclassifications/2011ruralurban-classification (accessed 1 November 2021).

97. Rokni K, Musa TA. Normalized difference vegetation change index: a technique for detecting vegetation changes using Landsat imagery. *CATENA* 2019;**178**:59–63. https://doi.org/10.1016/j.catena.2019.03.007

98. Lilley B. The ordnance survey opendata initiative. *Cartogr J* 2011;**48**:179–82. https://doi.org/10.1179/000870411X13044121958821

99. Liao Y, Zhou Q, Jing X. A comparison of global and regional open datasets for urban greenspace mapping. *Urban For Urban Green* 2021;**62**:127132. https://doi.org/10.1016/j.ufug.2021.127132

100. Welsh Assembly Government. *Technical Advice Note 16: Sport, Recreation and Open Space*. Cardiff: Welsh Assembly Government; 2009.

101. Bennett J. *OpenStreetMap: Be Your Own Cartographer*. Birmingham: Packt Publishing; 2010.

102. Helme M, Brown Z, Perez-Dominguez R. *National Survey for Wales 2018–19: Technical Report*. Cardiff: Welsh Government; 2019.

103. John A, McGregor J, Fone D, Dunstan F, Cornish R, Lyons RA, *et al.* Case-finding for common mental disorders of anxiety and depression in primary care: an external validation of routinely collected data. *BMC Med Inform Decis Mak* 2016;**16**:35. https://doi.org/10.1186/s12911-016-0274-7

104. Rodgers SE, Bailey R, Johnson R, Poortinga W, Smith R, Berridge D, *et al.* Health Impact, and Economic Value, of Meeting Housing Quality Standards: A Retrospective Longitudinal Data Linkage Study. *Public Health Res* 2018;**6**(8):1–103.

105. Bebbington PE, Meltzer H, Brugha TS, Farrell M, Jenkins R, Ceresa C, *et al.* Unequal access and unmet need: neurotic disorders and the use of primary care services. *Psychol Med* 2000;**30**:1359–67. https://doi.org/10.1017/s0033291799002950

106. Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, *et al.* The Warwick–Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. *Health Qual Life Outcomes* 2007;**5**:63. https://doi.org/10.1186/1477-7525-5-63

107. Tinkler L, Hicks S. *Measuring Subjective Well-Being*. London: Office for National Statistics; 2011.

108. Fone D, Morgan J, Fry R, Rodgers S, Orford S, Farewell D, *et al.* Change in Alcohol Outlet Density and Alcohol-Related Harm to Population Health (CHALICE): A Comprehensive Record-Linked Database Study in Wales. *Public Health Res* 2016;**4**(3):1–184. https://doi.org/10.3310/phr04030

109. Rodgers SE, Bailey R, Johnson R, Berridge D, Poortinga W, Lannon S, *et al.* Emergency hospital admissions associated with a non-randomised housing intervention meeting national housing quality standards: a longitudinal data linkage study. *J Epidemiol Community Health* 2018;**72**:896–903. https://doi.org/10.1136/jech-2017-210370

110. *SeRP*. SeRP. n.d. URL: https://serp.ac.uk/ (Accessed 8 November 2021).

111. Jones KH, Ford DV, Thompson S, Lyons R. A profile of the SAIL databank on the UK secure research platform. *Int J Popul Data Sci* 2019;**4**(2):1134. https://doi.org/10.23889/ijpds.v4i2.1134

112. Thayer D, Rees A, Kennedy J, Collins H, Harris D, Halcox J, *et al.* Measuring follow-up time in routinely-collected health datasets: challenges and solutions. *PLOS One* 2020;**15**:e0228545. https://doi.org/10.1371/journal.pone.0228545

113. Office for National Statistics. *2011 Census: Key Statistics for Wales, March 2011*. London: ONS; 2011. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/2011censuskeystatisticsforwales/2012-12-11 (accessed 1 November 2021).

114. Van den Bosch M. *Urban Green Spaces and Health – A Review of Evidence*. Copenhagen: World Health Organization Regional Office for Europe; 2016.

115. Labib SM, Lindley S, Huck JJ. Spatial dimensions of the influence of urban green-blue spaces on human health: a systematic review. *Environ Res* 2020;**180**:108869. https://doi.org/10.1016/j.envres.2019.108869

116. Institute for Leisure and Amenity Management. *Open Space Terminology: ILAM Fact Sheet 00/99 1999*. Reading: ILAM; 1999.

117. Swanwick C, Dunnett N, Woolley H. Nature, role and value of green space in towns and cities: an overview. *Built Environ (1978-)* 2003;**29**:94–106.

118. Kimpton A. A spatial analytic approach for classifying greenspace and comparing greenspace social equity. *Appl Geogr* 2017;**82**:129–42. https://doi.org/10.1016/j.apgeog.2017.03.016

119. Barbosa O, Tratalos JA, Armsworth PR, Davies RG, Fuller RA, Johnson P, *et al.* Who benefits from access to green space? A case study from Sheffield, UK. *Landsc Urban Plann* 2007;**83**:187–95. https://doi.org/10.1016/j.landurbplan.2007.04.004

120. Bell S, Montarzino A, Travlou P. Mapping research priorities for green and public urban space in the UK. *Urban For Urban Green* 2007;**6**:103–15. https://doi.org/10.1016/j.ufug.2007.03.005

121. Cvejić R, Eler K, Pintar M, Železnikar Š, Haase D, Kabisch N, Strohbach M. *A Typology of Urban Green Spaces, Ecosystem Services Provisioning Services and Demands*. Report D3.1. Brussels: European Union Seventh Framework Programme; 2015. URL: https://assets.centralparknyc.org/pdfs/institute/p2p-upelp/1.004_Greensurge_A+Typology+of+Urban+Green+Spaces.pdf (accessed 19 April 2023).

122. Hunter AJ, Luck GW. Defining and measuring the social-ecological quality of urban greens-pace: a semi-systematic review. *Urban Ecosyst* 2015;**18**:1139–63. https://doi.org/10.1007/s11252-015-0456-6

123. Natural England. *Nature Nearby: Accessible Natural Greenspace Guidance NE265*. London: Natural England; 2010. URL: https://webarchive.nationalarchives.gov.uk/20140605145320/http://publications.naturalengland.org.uk/publication/40004?category=47004 (accessed 11 February 2021).

124. Office for National Statistics. *UK Natural Capital: Interim Review and Revised 2020 Roadmap*. London: ONS; 2018. URL: https://www.ons.gov.uk/economy/environmentalaccounts/methodologies/uknaturalcapitalinterimreviewandrevised2020roadmap (accessed 16 February 2021).

125. *UK National Ecosystem Assessment*. 2012. URL: http://uknea.unep-wcmc.org/About/NEAReportStructure/tabid/62/Default.aspx (accessed 16 February 2021).

126. Scottish Government. *Planning and Open Space*. Planning Advice Note 65. Edinburgh: Scottish Government; 2008. URL: https://www.gov.scot/publications/planning-advice-note-pan-65-planning-open-space/pages/5/ (Accessed 16 February 2021).

127. Department for Communities and Local Government. *Planning Policy Guidance 17: Planning for Open Space, Sport and Recreation*. London: Department for Communities and Local Government; 2002. URL: http://www.greeninfrastructurenw.co.uk/climatechange/doc.php?docID=99 (accessed 19 April 2023).

128. Griffiths LA. *Future Landscapes: Delivering for Wales. The Review of Areas of Outstanding Natural Beauty and National Parks in Wales*. Cardiff: Department for Environment and Rural Affairs; 2017.

129. National Assembly for Wales. *Planning (Wales) Act 2015*. URL: https://www.legislation.gov.uk/anaw/2015/4/contents/enacted (accessed 19 April 2023).

130. Natural England. *Accessible Natural Greenspace Standard (ANGSt)*. London: Natural England; 2013. URL: https://webarchive.nationalarchives.gov.uk/20140605111422/http://www.naturalengland.org.uk/regions/east_of_england/ourwork/gi/accessiblenaturalgreenspace-standardangst.aspx (accessed 16 February 2021).

131. Fields in Trust. *Guidance for Outdoor Sport and Play: Beyond the Six Acre Standard, Wales*. Cardiff: Fields in Trust Cymru; 2017.

132. Clark RN, Stankey GH. *The Recreation Opportunity Spectrum: A Framework for Planning, Management, And Research. General Technical Report PNW-98*. Portland, OR: Pacific Northwest Region Forest Service; 1979. https://doi.org/10.1177/004728758001900244

133. World Health Organization Regional Office for Europe. *Urban Green Spaces and Health: A Review of Evidence*. Copenhagen: WHO Regional Office for Europe; 2016.

134. Planning Portal. *Planning Portal Glossary*. PortalPlanQuest; 2023. URL: https://www.planning-portal.co.uk/directory/4/glossary (accessed 6 July 2021).

135. Tratalos JA, Haines-Young R, Potschin M, Fish R, Church A. Cultural ecosystem services in the UK: lessons on designing indicators to inform management and policy. *Ecol Indic* 2016;**61**:63–73. https://doi.org/10.1016/j.ecolind.2015.03.040

136. Dadvand P, Wright J, Martinez D, Basagaña X, McEachan RRC, Cirach M, *et al.* Inequality, green spaces, and pregnant women: roles of ethnicity and individual and neighbourhood socio-economic status. *Environ Int* 2014;**71**:101–8. https://doi.org/10.1016/j.envint.2014.06.010

137. Sarkar C. Residential greenness and adiposity: findings from the UK Biobank. *Environ Int* 2017;**106**:1–10. https://doi.org/10.1016/j.envint.2017.05.016

138. Dzhambov A, Hartig T, Markevych I, Tilov B, Dimitrova D. Urban residential greenspace and mental health in youth: different approaches to testing multiple pathways yield different conclusions. *Environ Res* 2018;**160**:47–59. https://doi.org/10.1016/j.envres.2017.09.015

139. White MP, Alcock I, Wheeler BW, Depledge MH. Coastal proximity, health and well-being: results from a longitudinal panel survey. *Health Place* 2013;**23**:97–103. https://doi.org/10.1016/j.healthplace.2013.05.006

140. Elliott LR, White MP, Grellier J, Rees SE, Waters RD, Fleming LE. Recreational visits to marine and coastal environments in England: where, what, who, why, and when? *Marine Policy* 2018;**97**:305–14. https://doi.org/10.1016/j.marpol.2018.03.013

141. Dennis M, James P. Evaluating the relative influence on population health of domestic gardens and green space along a rural-urban gradient. *Landsc Urban Plann* 2017;**157**:343–51. https://doi.org/10.1016/j.landurbplan.2016.08.009

142. van den Bosch M, Sang O. Urban natural environments as nature-based solutions for improved public health – a systematic review of reviews. *Environ Res* 2017;**158**:373–84. https://doi.org/10.1016/j.envres.2017.05.040

143. US Geological Survey. *EarthExplorer: Home*. n.d. URL: https://earthexplorer.usgs.gov (accessed 2 April 2020).

144. *QGIS Python Plugins Repository*. n.d. URL: https://plugins.qgis.org/plugins/SemiAutomaticClassificationPlugin (accessed 2 April 2020).

145. Young NE, Anderson RS, Chignell SM, Vorster AG, Lawrence R, Evangelista PH. A survival guide to Landsat preprocessing. *Ecology* 2017;**98**:920–32. https://doi.org/10.1002/ecy.1730

146. GRASS Development Team. *GRASS GIS Manual i.vi*. Open Source Geospatial Foundation; 2022. URL: https://grass.osgeo.org/grass78/manuals/i.vi.html (accessed 2 April 2020).

147. *QGIS Python Plugins Repository: Cloud Masking*. n.d. URL: https://plugins.qgis.org/plugins/CloudMasking (accessed 19 April 2023).

148. Ordnance Survey. *OS MasterMap Topography Layer*. 2017. URL: https://www.ordnancesurvey.co.uk/business-and-government/products/topography-layer.html (accessed 15 November 2017).

149. Ordnance Survey. OS Greenspace. Southampton: OS; 2023.

150. DataMap Wales. *Lle A Geo-Portal for Wales*. 2023. URL: https://datamap.gov.wales/info/lle-decommissioned (accessed 19 April 2023).

151. DataMap Wales. *What We Do*. 2023. URL: https://datamap.gov.wales/info/what-we-do (accessed 19 April 2023).

152. OpenStreetMap. *Wales*. 2018. URL: http://download.geofabrik.de/europe/great-britain/wales.html (accessed 2 February 2017).

153. Biernacka M, Kronenberg J. Classification of institutional barriers affecting the availability, accessibility and attractiveness of urban green spaces. *Urban Forest Urban Green* 2018;**36**:22–33. https://doi.org/10.1016/j.ufug.2018.09.007

154. Hitchings R. Studying the preoccupations that prevent people from going into green space. *Landsc Urban Plann* 2013;**118**:98–102. https://doi.org/10.1016/j.landurbplan.2012.09.006

155. Natural England. *Monitor of Engagement with the Natural Environment (MENE)*. London: Natural England; 2014. URL: https://www.gov.uk/government/collections/monitor-of-engagement-with-the-natural-environment-survey-purpose-and-results (accessed 2 July 2021).

156. White MP, Pahl S, Wheeler BW, Depledge MH, Fleming LE. Natural environments and subjective wellbeing: different types of exposure are associated with different aspects of wellbeing. *Health Place* 2017;**45**:77–84. https://doi.org/10.1016/j.healthplace.2017.03.008

157. SAIL Databank. *Welsh Demographic Service Dataset*. SAIL; 2023. URL: https://web.www. healthdatagateway.org/dataset/8a8a5e90-b0c6-4839-bcd2-c69e6e8dca6d (accessed 19 April 2023).

158. Johnson RD, Griffiths LJ, Hollinghurst JP, Akbari A, Lee A, Thompson DA, *et al.* Deriving household composition using population-scale electronic health record data: a reproducible methodology. *PLOS One* 2021;**16**:e0248195. https://doi.org/10.1371/journal.pone.0248195

159. Rodgers SE, Lyons RA, Dsilva R, Jones KH, Brooks CJ, Ford DV, *et al.* Residential Anonymous Linking Fields (RALFs): a novel information infrastructure to study the interaction between the environment and individuals' health. *J Public Health* 2009;**31**:582–8. https://doi.org/10.1093/pubmed/fdp041

160. Helbich M, O'Connor RC, Nieuwenhuijsen M, Hagedoorn P. Greenery exposure and suicide mortality later in life: a longitudinal register-based case-control study. *Environ Int* 2020;**143**:105982. https://doi.org/10.1016/J.ENVINT.2020.105982

161. Ordnance Survey. *OS MasterMap Revision Policy*. 2023. URL: https://www.ordnancesurvey. co.uk/business-government/tools-support/mastermap-topography-support/revision-policy (accessed 19 April 2023).

162. Triguero-Mas M, Dadvand P, Cirach M, Martínez D, Medina A, Mompart A, *et al.* Natural outdoor environments and mental and physical health: relationships and mechanisms. *Environ Int* 2015;**77**:35–41. https://doi.org/10.1016/j.envint.2015.01.012

163. *Measuring Vegetation (NDVI & EVI)*. NASA Earth Observatory; 2000. URL: https://earthobserva-tory.nasa.gov/features/MeasuringVegetation (Accessed 8 November 2021).

164. Mental Health Foundation. *Fundamental Facts About Mental Health in Wales 2016*. London: Mental Health Foundation; 2016. URL: https://www.mentalhealth.org.uk/explore-mental-health/publications/fundamental-facts-about-mental-health-2016 (accessed 19 April 2023).

165. Engemann K, Pedersen CB, Arge L, Tsirogiannis C, Mortensen PB, Svenning J-C. Residential green space in childhood is associated with lower risk of psychiatric disorders from adoles-cence into adulthood. *PNAS* 2019;**116**:5188–93. https://doi.org/10.1073/pnas.1807504116

166. Gascon M, Cirach M, Martínez D, Dadvand P, Valentín A, Plasència A, *et al.* Normalized difference vegetation index (NDVI) as a marker of surrounding greenness in epidemiological studies: the case of Barcelona city. *Urban Forest Urban Green* 2016;**19**:88–94. https://doi.org/10.1016/J.UFUG.2016.07.001

167. Dadvand P, Bartoll X, Basagaña X, Dalmau-Bueno A, Martinez D, Ambros A, *et al.* Green spaces and general health: roles of mental health status, social support, and physical activity. *Environ Int* 2016;**91**:161–7. https://doi.org/10.1016/j.envint.2016.02.029

168. Sugiyama T, Leslie E, Giles-Corti B, Owen N. Associations of neighbourhood greenness with physical and mental health: do walking, social coherence and local social interaction explain the relationships? *J Epidemiol Community Health* 2008;**62**:e9. https://doi.org/10.1136/JECH.2007.064287

169. White MP, Pahl S, Ashbullby K, Herbert S, Depledge MH. Feelings of restoration from recent nature visits. *J Environ Psychol* 2013;**35**:40–51. https://doi.org/10.1016/j.jenvp.2013.04.002

170. World Health Organization Regional Office for Europe. *Urban Green Spaces: A Brief for Action*. Copenhagen: World Health Organization Regional Office for Europe; 2017.

171. Hunter RF, Cleland C, Cleary A, Droomers M, Wheeler BW, Sinnett D, *et al.* Environmental, health, wellbeing, social and equity effects of urban green space interventions: a meta-narra-tive evidence synthesis. *Environ Int* 2019;**130**. https://doi.org/10.1016/J.ENVINT.2019.104923

172. Rigolon A, Browning MHEM, McAnirlin O, Yoon H. Green space and health equity: a systematic review on the potential of green space to reduce health disparities. *Int J Environ Res Public Health* 2021;**18**:2563. https://doi.org/10.3390/ijerph18052563

173. Mitchell R, Astell-Burt T, Richardson EA. A comparison of green space indicators for epidemiological research. *J Epidemiol Community Health* 2011;**65**:853–8. https://doi.org/10.1136/jech.2010.119172

174. Mitchell RJ, Popham F. Effect of exposure to natural environment on health inequalities: an observational population study. *Lancet* 2008;**372**:1655–60. https://doi.org/10.1016/S0140-6736(08)61689-X

175. Roberts M, Irvine KN, McVittie A. Associations between greenspace and mental health prescription rates in urban areas. *Urban Forest Urban Green* 2021;**64**:127301. https://doi.org/10.1016/J.UFUG.2021.127301

176. Nutsford D, Pearson AL, Kingham S. An ecological study investigating the association between access to urban green space and mental health. *Public Health* 2013;**127**:1005–11. https://doi.org/10.1016/j.puhe.2013.08.016

177. Wood L, Hooper P, Foster S, Bull F. Public green spaces and positive mental health – investigating the relationship between access, quantity and types of parks and mental wellbeing. *Health Place* 2017;**48**:63–71. https://doi.org/10.1016/j.healthplace.2017.09.002

178. Galea S, Ahern J, Rudenstine S, Wallace Z, Vlahov D. Urban built environment and depression: a multilevel analysis. *J Epidemiol Community Health* 2005;**59**:822–7. https://doi.org/10.1136/JECH.2005.033084

179. Wu L, Kim SK. Health outcomes of urban green space in China: evidence from Beijing. *Sustain Cities Soc* 2021;**65**:102604. https://doi.org/10.1016/J.SCS.2020.102604

180. Kabisch N, Qureshi S, Haase D. Human–environment interactions in urban green spaces — a systematic review of contemporary issues and prospects for future research. *Environ Impact Assess Rev* 2015;**50**:25–34. https://doi.org/10.1016/J.EIAR.2014.08.007

181. Kelly D, Steiner A, Mazzei M, Baker R. Filling a void? The role of social enterprise in addressing social isolation and loneliness in rural communities. *J Rural Stud* 2019;**70**:225–36. https://doi.org/10.1016/J.JRURSTUD.2019.01.024

182. Henning-Smith C, Ecklund A, Lahr M, Evenson A, Moscovice I, Kozhimannil K. *Key Informant Perspectives on Rural Isolation and Loneliness*. Policy Brief. Minneapolis, MN: Rural Health Research & Policy Centers; 2018.

183. Jordan H, Roderick P, Martin D, Barnett S. Distance, rurality and the need for care: access to health services in South West England. *Int J Health Geogr* 2004;**3**:21. https://doi.org/10.1186/1476-072X-3-21

184. Mizen A, Fry R, Grinnell D, Rodgers S E. Quantifying the error associated with alternative GIS-based techniques to measure access to health care services. *Health Services Res* 2015;**2**:439–45. https://doi.org/10.1136/jech.2005.043281

185. Fotheringham, D AS, Wong WS. The modifiable areal unit problem in multivariate statistical analysis. *Environ Plann A* 1991;**23**:1025–44. https://doi.org/10.1068/a231025

186. Hobbs M, Tomintz M, McCarthy J, Marek L, Vannier C, Campbell M, *et al.* Obesity risk in women of childbearing age in New Zealand: a nationally representative cross-sectional study. *Int J Public Health* 2019;**64**:625–35. https://doi.org/10.1007/S00038-019-01239-8

187. Zhang L, Zhou S, Kwan M-P, Chen F, Lin R. Impacts of individual daily greenspace exposure on health based on individual activity space and structural equation modeling. *Int J Environ Res Public Health* 2018;**15**:2323. https://doi.org/10.3390/IJERPH15102323

124

188. Anguelovski I, Connolly JJT, Masip L, Pearsall H. Assessing green gentrification in historically disenfranchised neighborhoods: a longitudinal and spatial analysis of Barcelona. *Urban Geogr* 2018;**39**:458–91. https://doi.org/10.1080/02723638.2017.1349987

189. Clarke GM, Conti S, Wolters AT, Steventon A. Evaluating the impact of healthcare interventions using routine data. *BMJ* 2019;**365**:l12239. https://doi.org/10.1136/BMJ.L2239

190. Althubaiti A. Information bias in health research: definition, pitfalls, and adjustment methods. *J Multidiscip Healthc* 2016;**9**:211. https://doi.org/10.2147/JMDH.S104807

191. Song H, Lane KJ, Kim H, Kim H, Byun G, Le M, *et al.* Association between urban greenness and depressive symptoms: evaluation of greenness using various indicators. *Int J Environ Res Public Health* 2019;**16**:173. https://doi.org/10.3390/ijerph16020173

192. Hystad P, Davies HW, Frank L, Loon JV, Gehring U, Tamburic L, *et al.* Residential greenness and birth outcomes: evaluating the influence of spatially correlated built-environment factors. *Environ Health Perspect* 2014;**122**:1095–102. https://doi.org/10.1289/ehp.1308049

193. Wang P, Meng YY, Lam V, Ponce N. Green space and serious psychological distress among adults and teens: a population-based study in California. *Health Place* 2019;**56**:184–90. https://doi.org/10.1016/j.healthplace.2019.02.002

194. Persson A, Pyko A, Lind T, Bellander T, Östenson C-G, Pershagen G, *et al.* Urban residential greenness and adiposity: a cohort study in Stockholm County. *Environ Int* 2018;**121**:832–41. https://doi.org/10.1016/j.envint.2018.10.009

195. Charlson ME, Carrozzino D, Guidi J, Patierno C. Charlson Comorbidity Index: a critical review of clinimetric properties. *Psychother Psychosom* 2022:**91**:8–35.

196. Cook JA, Ranstam J. Statistical models and confounding adjustment. *BJS (Br J Surg)* 2017;**104**:786–7. https://doi.org/10.1002/bjs.10245

197. NHS Digital. *Investment in General Practice 2013–14 to 2017–18 England, Wales, Northern Ireland and Scotland*. London: NHS Digital; 2018. URL: https://digital.nhs.uk/data-and-information/publications/statistical/investment-in-general-practice/2013-14-to-2017-18-england-wales-northern-ireland-and-scotland (accessed 19 April 2023).

198. Reklaitiene R, Grazuleviciene R, Dedele A, Virviciute D, Vensloviene J, Tamosiunas A, *et al.* The relationship of green space, depressive symptoms and perceived general health in urban population. *Scand J Public Health* 2014;**42**:669–76. https://doi.org/10.1177/1403494814544494

199. Anguelovski I. From toxic sites to parks as (Green) LULUs? New challenges of inequity, privilege, gentrification, and exclusion for urban environmental justice. *J Plan Lit* 2016;**31**:23–36. https://doi.org/10.1177/0885412215610491

200. Schnake-Mahl AS, Jahn JL, Subramanian SV, Waters MC, Arcaya M. Gentrification, neighborhood change, and population health: a systematic review. *J Urban Health* 2020;**97**:1–25. https://doi.org/10.1007/s11524-019-00400-1

201. Elhag M, Boteva S, Al-Amri N. Forest cover assessment using remote-sensing techniques in Crete Island, Greece. *Open Geosci* 2021;**13**:345–58. https://doi.org/10.1515/geo-2020-0235

202. Boone-Heinonen J, Guilkey DK, Evenson KR, Gordon-Larsen P. Residential self-selection bias in the estimation of built environment effects on physical activity between adolescence and young adulthood. *Int J Behav Nutr Phys Act* 2010;**7**:70. https://doi.org/10.1186/1479-5868-7-70

203. Burns EM, Rigby E, Mamidanna R, Bottle A, Aylin P, Ziprin P, *et al.* Systematic review of discharge coding accuracy. *J Public Health (Oxf)* 2012;**34**:138–48. https://doi.org/10.1093/pubmed/fdr054

204. Larvin H, Peckham E, Prady SL. Case-finding for common mental disorders in primary care using routinely collected data: a systematic review. *Soc Psychiatry Psychiatr Epidemiol* 2019;**54**:1161–75. https://doi.org/10.1007/s00127-019-01744-4

205. Linton M-J, Dieppe P, Medina-Lara A. Review of 99 self-report measures for assessing well-being in adults: exploring dimensions of well-being and developments over time. *BMJ Open* 2016;**6**:e010641. https://doi.org/10.1136/bmjopen-2015-010641

206. World Health Organization. *Mental Health Action Plan 2013–2020*. Geneva: WHO; 2013. https://www.who.int/publications/i/item/9789241506021 (accessed 19 April 2023).

207. Schotanus-Dijkstra M, ten Have M, Lamers SMA, de Graaf R, Bohlmeijer ET. The longitudinal relationship between flourishing mental health and incident mood, anxiety and substance use disorders. *Eur J Public Health* 2017;**27**:563–8. https://doi.org/10.1093/eurpub/ckw202

208. Rehm J, Shield KD. Global burden of disease and the impact of mental and addictive disorders. *Curr Psychiatry Rep* 2019;**21**:10. https://doi.org/10.1007/s11920-019-0997-0

209. World Health Organization. *Depression and Other Common Mental Disorders: Global Health Estimates*. Geneva: WHO; 2017.

210. Tillmann S, Tobin D, Avison W, Gilliland J. Mental health benefits of interactions with nature in children and teenagers: a systematic review. *J Epidemiol Community Health* 2018;**72**:958–66. https://doi.org/10.1136/jech-2018-210436

211. Gascon M, Sánchez-Benavides G, Dadvand P, Martínez D, Gramunt N, Gotsens X, *et al.* Long-term exposure to residential green and blue spaces and anxiety and depression in adults: a cross-sectional study. *Environ Res* 2018;**162**:231–9. https://doi.org/10.1016/j.envres.2018.01.012

212. Tester-Jones M, White MP, Elliott LR, Weinstein N, Grellier J, Economou T, *et al.* Results from an 18 country cross-sectional study examining experiences of nature for people with common mental health disorders. *Sci Rep* 2020;**10**:19408. https://doi.org/10.1038/s41598-020-75825-9

213. Elliott LR, White MP, Grellier J, Garrett JK, Cirach M, Wheeler BW, *et al.* Research Note: residential distance and recreational visits to coastal and inland blue spaces in eighteen countries. *Landsc Urban Plann* 2020;**198**:103800. https://doi.org/10.1016/j.landurbplan.2020.103800

214. Schipperijn J, Ekholm O, Stigsdotter UK, Toftager M, Bentsen P, Kamper-Jørgensen F, *et al.* Factors influencing the use of green space: results from a Danish national representative survey. *Landsc Urban Plann* 2010;**95**:130–7. https://doi.org/10.1016/j.landurbplan.2009.12.010

215. White MP, Wheeler BW, Herbert S, Alcock I, Depledge MH. Coastal proximity and physical activity: is the coast an under-appreciated public health resource? *Prev Med* 2014;**69**:135–40. https://doi.org/10.1016/j.ypmed.2014.09.016

216. Börger T, Campbell D, White MP, Elliott LR, Fleming LE, Garrett JK, *et al.* The value of blue-space recreation and perceived water quality across Europe: a contingent behaviour study. *Sci Total Environ* 2021;**771**:145597. https://doi.org/10.1016/j.scitotenv.2021.145597

217. van den Berg M, van Poppel M, Smith G, Triguero-Mas M, Andrusaityte S, van Kamp I, *et al.* Does time spent on visits to green space mediate the associations between the level of residential greenness and mental health? *Urban Forest Urban Green* 2017;**25**:94–102. https://doi.org/10.1016/j.ufug.2017.04.010

218. Bojorquez I, Ojeda-Revah L. Urban public parks and mental health in adult women: mediating and moderating factors. *Int J Soc Psychiatry* 2018;**64**:637–46. https://doi.org/10.1177/0020764018795198

219. Orstad SL, Szuhany K, Tamura K, Thorpe LE, Jay M. Park proximity and use for physical activity among urban residents: associations with mental health. *Int J Environ Res Public Health* 2020;**17**:4885. https://doi.org/10.3390/ijerph17134885

220. Bellis MA, Lowey H, Hughes K, Deacon L, Stansfield J, Perkins C. Variations in risk and protective factors for life satisfaction and mental wellbeing with deprivation: a cross-sectional study. *BMC Public Health* 2012;**12**:492. https://doi.org/10.1186/1471-2458-12-492

221. Jones L, Holland RA, Ball J, Sykes T, Taylor G, Ingwall-King L, *et al.* A place-based participatory mapping approach for assessing cultural ecosystem services in urban green space. *People Nat* 2020;**2**:123–37. https://doi.org/10.1002/pan3.10057

222. Marmot M. Health equity in England: the Marmot review 10 years on. *BMJ* 2020;**368**:m693. https://doi.org/10.1136/bmj.m693

223. Ministry of Housing, Communities and Local Government. *English Indices of Deprivation 2019*. London: Ministry of Housing, Communities and Local Government; 2020. URL: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019 (accessed 19 April 2023).

224. Public Health England. *Improving Access to Greenspace: A New Review for 2020*. London: Public Health England; 2020.

225. Hunter RF, Cleary A, Braubach M. Environmental, health and equity effects of urban green space interventions. In Marselle M, Stadler J, Korn H, Irvine K, Bonn A, editors. *Biodiversity and Health in the Face of Climate Change*. Cham: Springer; 2019. pp. 381–409.

226. Welsh Government. *National Survey for Wales 2016–17 Questionnaire*. Cardiff: Welsh Government; 2016. URL: https://www.gov.wales/national-survey-wales-questionnaires (accessed 19 April 2023).

227. Welsh Government. *National Survey for Wales 2018–19 Questionnaire*. Cardiff: Welsh Government; 2018. URL: https://gov.wales/national-survey-wales-questionnaires (accessed 19 April 2023).

228. Aumeyr M, Brown Z, Doherty R, Fallows A, Pegg T, Perez-dominguez R, *et al. National Survey for Wales 2016–17 Technical Report (Updated)*. Cardiff: Welsh Government; 2017. URL: http://doc.ukdataservice.ac.uk/doc/8301/mrdoc/pdf/8301_171018-national-survey-wales-2016-17-technical-report-en.pdf (accessed 19 April 2023).

229. McManus S, Bebbington PE, Jenkins R, Brugha T. *Mental Health and Wellbeing in England: The Adult Psychiatric Morbidity Survey 2014*. Leeds: NHS Digital; 2016. URL: https://digital.nhs.uk/data-and-information/publications/statistical/adult-psychiatric-morbidity-survey/adult-psychiatric-morbidity-survey-survey-of-mental-health-and-wellbeing-england-2014 (accessed 19 April 2023).

230. Maheswaran H, Weich S, Powell J, Stewart-Brown S. Evaluating the responsiveness of the Warwick Edinburgh Mental Well-Being Scale (WEMWBS): group and individual level analysis. *Health Qual Life Outcomes* 2012;**10**:156. https://doi.org/10.1186/1477-7525-10-156

231. Mitchell R. Is physical activity in natural environments better for mental health than physical activity in other environments? *Soc Sci Med* 2013;**91**:130–4. https://doi.org/10.1016/j.socscimed.2012.04.012

232. VanderWeele TJ, Robins JM. Four types of effect modification a classification based on directed acyclic graphs. *Epidemiology* 2007;**18**:561–8. https://doi.org/10.1097/EDE.0b013e318127181b

233. Corraini P, Olsen M, Pedersen L, Dekkers OM, Vandenbroucke JP. Effect modification, interaction and mediation: an overview of theoretical insights for clinical investigators. *Clin Epidemiol* 2017;**9**:331–8. https://doi.org/10.2147/CLEP.S129728

234. van Stralen KJ, Dekker FW, Zoccali C, Jager K. Confounding. *NEC* 2010;**116**:c143–7. https://doi.org/10.1159/000315883

235. World Health Organization. *Global Recommendations on Physical Activity for Health*. Geneva: WHO; 2010. URL: https://www.who.int/publications/i/item/9789241599979 (accessed 19 April 2023).

236. Department of Health and Social Care, Welsh Government, Department of Health Northern Ireland and the Scottish Government. *UK Chief Medical Officers' Physical Activity Guidelines*. London: DHSC; 2019. URL: https://www.gov.uk/government/publications/physical-activity-guidelines-uk-chief-medical-officers-report (accessed 19 April 2023).

237. Hafferty C. *What Factors Are Linked to People Living in Households That Are in Material Deprivation?*. Cardiff: Welsh Government; 2020. URL: https://gov.wales/sites/default/files/statistics-and-research/2020-03/what-factors-are-linked-to-people-living-in-households-that-are-in-material-deprivation.pdf (accessed 12 August 2021).

238. Wood SN. *Generalized Additive Models: An Introduction with R*. 2nd edn. Boca Raton, FL: CRC Press; 2017.

239. Baron RM, Kenny DA. The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986;**51**:1173–82. https://doi.org/10.1037/0022-3514.51.6.1173

240. Lüdecke D, Makowski D, Ben-Shachar MS, Patil I, Højsgaard S. *Package 'Parameters': Processing of Model Paramaters. Version 0.12.0*. 2021. URL: https://cran.r-project.org/web/packages/parameters/parameters.pdf (accessed 19 April 2023).

241. World Health Organization Regional Office for Europe. *Urban Green Spaces and Health: A Review of Evidence*. Copenhagen; WHO Regional Office for Europe; 2016.

242. Crouse DL, Pinault L, Christidis T, Lavigne E, Thomson EM, Villeneuve PJ. Residential greenness and indicators of stress and mental well-being in a Canadian national-level survey. *Environ Res* 2021;**192**:110267. https://doi.org/10.1016/j.envres.2020.110267

243. Kruize H, van Kamp I, van den Berg M, van Kempen E, Wendel-Vos W, Ruijsbroek A, *et al*. Exploring mechanisms underlying the relationship between the natural outdoor environment and health and well-being – results from the PHENOTYPE project. *Environ Int* 2020;**134**:105173. https://doi.org/10.1016/j.envint.2019.105173

244. Frumkin H, Bratman GN, Breslow SJ, Cochran B, Kahn PH, Lawler JJ, *et al*. Nature contact and human health: a research agenda. *Environ Health Perspect* 2017;**125**:1–18. https://doi.org/10.1289/EHP1663

245. van den Berg M, van Poppel M, van Kamp I, Ruijsbroek A, Triguero-Mas M, Gidlow C, *et al*. Do physical activity, social cohesion, and loneliness mediate the association between time spent visiting green space and mental health? *Environ Behav* 2017;**51**:144–66.

246. White MP, Elliott LR, Wheeler BW, Fleming LE. Neighbourhood greenspace is related to physical activity in England, but only for dog owners. *Landsc Urban Plann* 2018;**174**:18–23. https://doi.org/10.1016/j.landurbplan.2018.01.004

247. Natural England. *Monitor of Engagement with the Natural Environment: Headline Report and Technical Reports 2018 to 2019*. London: Natural England; 2019. URL: https://www.gov.uk/government/statistics/monitor-of-engagement-with-the-natural-environment-headline-report-and-technical-reports-2018-to-2019 (accessed 14 August 2021).

248. Daniel RM, De Stavola BL, Cousens SN. Gformula: estimating causal effects in the presence of time-varying confounding or mediation using the G-computation formula. *Stata J* 2011;**11**:479–517. https://doi.org/10.1177/1536867X1201100401

249. Mell I. The impact of austerity on funding green infrastructure: a DPSIR evaluation of the Liverpool Green & Open Space Review (LG&OSR), UK. *Land Use Policy* 2020;**91**:104284. https://doi.org/10.1016/j.landusepol.2019.104284

250. Mell I. Financing the future of green infrastructure planning: alternatives and opportunities in the UK. *Landscape Res* 2018;**43**:751–68. https://doi.org/10.1080/01426397.2017.1390079

251. Rex B, Campbell P. The impact of austerity measures on local government funding for culture in England. *Cult Trends* 2021;**0**:1–24. https://doi.org/10.1080/09548963.2021.1915096

252. Shanahan DF, Bush R, Gaston KJ, Lin BB, Dean J, Barber E, *et al.* Health benefits from nature experiences depend on dose. *Sci Rep* 2016;**6**:28551. https://doi.org/10.1038/srep28551

253. Garcia LMT, Diez Roux AV, Martins ACR, Yang Y, Florindo AA. Exploring the emergence and evolution of population patterns of leisure-time physical activity through agent-based modelling. *Int J Behav Nutr Phys Act* 2018;**15**:112. https://doi.org/10.1186/s12966-018-0750-9

254. Badham J, Chattoe-Brown E, Gilbert N, Chalabi Z, Kee F, Hunter R. Developing agent-based models of complex health behaviour. *Health Place* 2018;**54**:170–7. https://doi.org/10.1016/j.healthplace.2018.08.022

255. de Vocht F, Katikireddi SV, McQuire C, Tilling K, Hickman M, Craig P. Conceptualising natural and quasi experiments in public health. *BMC Med Res Methodol* 2021;**21**:32. https://doi.org/10.1186/s12874-021-01224-x

256. Craig P, Cooper C, Gunnell D, Haw S, Lawson K, Macintyre S, *et al.* Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. *J Epidemiol Community Health* 2012;**66**:1182–6. https://doi.org/10.1136/jech-2011-200375

257. Remler DK, Ryzin GGV. *Research Methods in Practice: Strategies for Description and Causation.* 2nd edn. Los Angeles, CA: SAGE; 2014.

258. Shadish W, Cook T, Campbell D. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* 2nd edn. Belmont, CA: Wadsworth; 2001.

259. Jerrett M, van den Bosch M. Nature exposure gets a boost from a cluster randomized trial on the mental health benefits of greening vacant lots. *JAMA Network Open* 2018;**1**:e180299. https://doi.org/10.1001/jamanetworkopen.2018.0299

260. Leatherdale ST. Natural experiment methodology for research: a review of how different methods can support real-world research. *Int J Soc Res Methodol* 2019;**22**:19–35. https://doi.org/10.1080/13645579.2018.1488449

261. Iyer HS, James P, Valeri L, Hart JE, Pernar CH, Mucci LA, *et al.* The association between neighborhood greenness and incidence of lethal prostate cancer. *Environ Epidemiol* 2020;**4**:e091. https://doi.org/10.1097/EE9.0000000000000091

262. Pearson AL, Pechal J, Lin Z, Benbow ME, Schmidt C, Mavoa S. Associations detected between measures of neighborhood environmental conditions and human microbiome diversity. *Sci Total Environ* 2020;**745**:141029. https://doi.org/10.1016/j.scitotenv.2020.141029

263. South EC, Hohl BC, Kondo MC, MacDonald JM, Branas CC. Effect of greening vacant land on mental health of community-dwelling adults: a cluster randomized trial. *JAMA Network Open* 2018;**1**:e180298. https://doi.org/10.1001/jamanetworkopen.2018.0298

264. Fuertes E, Markevych I, Thomas R, Boyd A, Granell R, Mahmoud O, *et al.* Residential greenspace and lung function up to 24 years of age: the ALSPAC birth cohort. *Environ Int* 2020;**140**:105749. https://doi.org/10.1016/j.envint.2020.105749

265. Geneshka M, Coventry P, Cruz J, Gilbody S. Relationship between green and blue spaces with mental and physical health: a systematic review of longitudinal observational studies. *Int J Environ Res Public Health* 2021;**18**:9010. https://doi.org/10.3390/ijerph18179010

266. Nuffield Trust. *Improving Access to Psychological Therapies (IAPT) Programme*. London: Nuffield Trust; 2018. URL: https://www.nuffieldtrust.org.uk/resource/improving-access-to-psychological-therapies-iapt-programme (accessed 19 October 2021).

267. Goldacre B, Morley J. *Better, Broader, Safer: Using Health Data for Research and Analysis. A Review Commissioned by the Secretary of State for Health and Social Care*. London: Department of Health and Social Care; 2022.

268. McDaid D, Park AL, Davidson G, John A, Knifton L, McDaid S, *et al. The Economic Case for Investing in the Prevention of Mental Health Conditions in the UK*. London: Mental Health Foundation; 2022. URL: https://www.mentalhealth.org.uk/sites/default/files/2022-06/MHF-Investing-in-Prevention-Full-Report.pdf (accessed 19 April 2023).

269. Cherrie MPC, Shortt NK, Mitchell RJ, Taylor AM, Redmond P, Thompson CW, *et al.* Green space and cognitive ageing: a retrospective life course analysis in the Lothian Birth Cohort 1936. *Soc Sci Med* 2018;**196**:56–65. https://doi.org/10.1016/j.socscimed.2017.10.038

270. HM Government. *Levelling Up the United Kingdom: Presented to Parliament by the Secretary of State for Levelling Up, Housing and Communities by Command of Her Majesty*. CP 604. London: HMSO; 2022. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1052064/Levelling_Up_White_Paper_HR.pdf (accessed 12 August 2022).

271. Harron KL, Doidge JC, Knight HE, Gilbert RE, Goldstein H, Cromwell DA, *et al.* A guide to evaluating linkage quality for the analysis of linked data. *Int J Epidemiol* 2017;**46**:1699–710. https://doi.org/10.1093/ije/dyx177.

272. Rutter H, Savona N, Glonti K, Bibby J, Cummins S, Finegood DT, *et al.* The need for a complex systems model of evidence for public health. *Lancet* 2017;**390**:2602–4. https://doi.org/10.1016/S0140-6736(17)31267-9

273. Geary RS, Wheeler B, Lovell R, Jepson R, Hunter R, Rodgers S. A call to action: improving urban green spaces to reduce health inequalities exacerbated by COVID-19. *Prev Med* 2021;**145**:106425. https://doi.org/10.1016/j.ypmed.2021.106425

274. Ward Thompson C, Oliveira ES, Tilley S, Elizalde A, Botha W, Briggs A, *et al.* Health impacts of environmental and social interventions designed to increase deprived communities' access to urban woodlands: a mixed-methods study. *Public Health Res* 2019;**7**(2). https://doi.org/10.3310/phr07020

275. Thompson DA, Geary RS, Rowney FM, Fry R, Watkins A, Wheeler BW, *et al.* Cohort Profile: the Green and Blue Spaces (GBS) and mental health in Wales e-cohort. *Int J Epidemiol* 2022;**51**(5):e285–94. https://doi.org/10.1093/ije/dyac080

276. *Google Scholar*. n.d. URL: https://scholar.google.com (accessed 18 December 2019).

277. *Web of Science* n.d. URL: https://www.webofknowledge.com (accessed 18 December 2019).

278. Almanza E, Jerrett M, Dunton G, Seto E, Ann Pentz M. A study of community design, greenness, and physical activity in children using satellite, GPS and accelerometer data. *Health Place* 2012;**18**:46–54. https://doi.org/10.1016/j.healthplace.2011.09.003

279. Klompmaker JO, Hoek G, Bloemsma LD, Gehring U, Strak M, Wijga AH, *et al.* Green space definition affects associations of green space with overweight and physical activity. *Environ Res* 2018;**160**:531–40. https://doi.org/10.1016/j.envres.2017.10.027

280. Kanholm IC. *Assessing Green Space as a Correlate of Physical Activity Among Twins*. Master in Public Health thesis. Washington, DC: University of Washington; 2013. URL: https://digital.lib. washington.edu/researchworks/handle/1773/24125?show=full (accessed 19 April 2023).

281. Rugel EJ, Henderson SB, Carpiano RM, Brauer M. Beyond the Normalized Difference Vegetation Index (NDVI): developing a Natural Space Index for population-level health research. *Environ Res* 2017;**159**:474–83. https://doi.org/10.1016/j.envres.2017.08.033

282. Rugel EJ, Carpiano RM, Henderson SB, Brauer M. Exposure to natural space, sense of community belonging, and adverse mental health outcomes across an urban region. *Environ Res* 2019;**171**:365–77. https://doi.org/10.1016/j.envres.2019.01.034

283. Rogerson M, Brown DK, Sandercock G, Wooller JJ, Barton J. A comparison of four typical green exercise environments and prediction of psychological health outcomes. *Perspect Public Health* 2016;**136**:171–80. https://doi.org/10.1177/1757913915589845

284. Grow HM, Saelens BE, Kerr J, Durant NH, Norman GJ, Sallis JF. Where are youth active? Roles of proximity, active transport, and built environment. *Med Sci Sports Exerc* 2008;**40**:2071–9. https://doi.org/10.1249/MSS.0b013e3181817baa

285. Dadvand P, Sunyer J, Basagaña X, Ballester F, Lertxundi A, Fernández-Somoano A, *et al.* Surrounding greenness and pregnancy outcomes in four Spanish birth cohorts. *Environ Health Perspect* 2012;**120**:1481–7. https://doi.org/10.1289/ehp.1205244

286. Ekkel ED, Vries S. Nearby green space and human health: evaluating accessibility metrics. *Landsc Urban Plann* 2017;**157**:214–20. https://doi.org/10.1016/j.landurbplan.2016.06.008

287. Apparicio P, Abdelmajid M, Riva M, Shearmur R. Comparing alternative approaches to measuring the geographical accessibility of urban health services: distance types and aggregation-error issues. *Int J Health Geogr* 2008;**7**:7. https://doi.org/10.1186/1476-072X-7-7

288. Fone D, Dunstan F, White J, Webster C, Rodgers S, Lee S, *et al.* Change in alcohol outlet density and alcohol-related harm to population health (CHALICE). *BMC Public Health* 2012;**12**:428. https://doi.org/10.1186/1471-2458-12-428

289. Mizen A. *Investigating the Impact of GIS Modelled Daily Exposures to the Retail Food Environment on Routinely Linked Child Health Data*. Swansea University; 2018.

Appendix 1　Residential anonymised linking field assignment conditions

The initial joining of a RALF to an individual's three-month period considers the following conditions:

- The RALF start date falls between the start of the quarter and end of the quarter. This captures a change in RALF during one quarter.
- The RALF start date is before the start of the quarter and the RALF end date is later than the end of the quarter. This captures RALFs that are linked throughout the whole quarter.
- The RALF start date is before period start date and the RALF end date is between the quarter start date and period end date and it is the first quarter the individual is in the study. This captures movement within the first quarter an individual enters the cohort.

When multiple RALFs are linked to an individual within a quarter, the conditions and quality ranks presented in the Table below are assigned; the highest ranked RALF was used as the most suitable single RALF within the quarter.

TABLE 22　Residential anonymised linking field assignment conditions

Condition	RALF assignment quality value
When one RALF is linked of high confidence	1
When one RALF is linked of intermediate confidence	2
When one RALF is linked of low confidence	3
When number of RALFs linked is > 1 and < 7: when RALF is of high confidence and RALF duration is the longest to 6th longest	4–9
When number of RALFs linked is > 1 and < 7: when RALF is of medium confidence and RALF is longest duration	10
When number of RALFs linked is > 1 and < 7: when RALF is of low confidence and RALF is longest duration	11
When no RALF is linked (RALF = NULL)	99
When number of RALFs linked is ≥ 7	99

Appendix 2 Typology tiers

TABLE 23 Description of typology tiers

Tier 1	Tier 2	Tier 3	Description
Amenity	Recreation space	Parks	Open to the public, areas of managed, open land; often with grass and formal flower-beds and/or trees
		Recreation areas	A purpose-built area used for playing games and social activities (e.g. bowling green, crazy golf, members' golf club)
		Sports pitches	Formal, maintained sports pitches (may be artificial grass)
		Play areas	A public area with swings, slides, climbing frames, etc.
	Private	Domestic gardens	A plot of ground adjacent to an addressed house
Functional	Productive	City farms	A farm within an urban area (often run educational programmes)
		Allotments	Small allocation of land, normally rented from a local authority
	Burial grounds	Cemeteries	An area where the dead are buried
		Churchyard	An area where the dead are buried, including graveyards and churchyards
	Institutional grounds	School grounds	A defined area belonging to a school premises
		Other grounds	An area of land that is associated with a property
	Gardens	Botanical gardens	An area of land containing plants of scientific interest and facilities for their care and cultivation
		Community garden	A single piece of land gardened collectively by a group of people
Semi-natural habitat	Wetland	Marsh	An area of land that is waterlogged throughout the year
	Woodland	Deciduous woodland	An area of ground covered with deciduous trees
		Coniferous woodland	An area of ground covered with coniferous trees
		Mixed woodland	An area of ground covered with diverse types of trees
	Other habitats	Moor/heath	An area of open, uncultivated, and usually high land with poor soil that is covered mainly with grass, heather, bilberry and sedges with very few trees or bushes
		Grassland	Areas of bracken and uncultivated grassland, often tufted with a 'hummocky' appearance, normally found on the higher parts of hills, mountains and down land
		Quarry	An area that has been dug for the purpose of extracting rock that is no longer active
		Meadow	An area of fertile grassland on the flood plain of a river

TABLE 23 Description of typology tiers *(continued)*

Tier 1	Tier 2	Tier 3	Description
Linear	Inland	River	Water flowing in a definite channel towards the sea, a lake or into another river
		Canal	An artificial watercourse for inland navigation
		Transport corridors	A purpose-built route for taking people from one place to another (road, rail, cycleways and walking)
	Coastal	Cliffs	A steep rock face, especially at the edge of the sea
		Beach	An area of sand, shingles or rocks beside inland or tidal water
		Marina	A small harbour or area of water where yachts and other pleasure craft can moor
		Docklands	An area in a harbour where ships are loaded/unloaded or repaired
		Estuary	The tidal mouth of a river where the channel broadens out at the coast
		Harbour	An area of naturally or artificially protected water on a coast where boats can anchor or moor
Enclosed blue feature	Inland	Lake	A large area of salt or fresh water which may be surrounded by land and is larger than a pond
		Reservoir	A natural or artificial lake or pond used for storing water which may be used for irrigation, water supply, hydroelectric power generation or flood control
		Pond	A small area of fresh water, often artificially created, which is smaller than a lake
		Outdoor swimming pool	Structure that contains water for swimming in
	Coastal	Lido	A public open-air swimming pool or bathing beach

Appendix 3 Definition of primary and secondary exposures

There is methodological heterogeneity in the evidence base on how we define exposure to GBS. We have calculated primary and secondary exposure measures to be linked with health data for the Welsh population from 2008 to 2019.

Aim

The main aim of this appendix is to document methodological choices that have been made in the creation of primary and secondary exposures for all households in Wales from 2008 to 2019.

Objectives

We will fulfil this aim by completing the following objectives:

1. evaluate how previous studies have defined and measured exposure to GBS
2. define primary exposure to GBS for this study.

Literature search

Eligibility criteria
We identified studies that investigated the relationship between mental health outcomes and green and blue spaces in adult populations. We included all study designs in our search but only included studies undertaken in developed countries. We included studies published before the end of 2019.

Information sources
We used online electronic databases to search for the studies: Google Scholar[276] and Web of Science.[277]

Search strategy

Study records
We used Mendeley (version 1.19.4) as our reference management software to manage the records and to remove duplicate studies. We kept records that met the following inclusion criteria:

i. no duplicates
ii. study population over 16 years of age
iii. primary outcome was a CMD such as anxiety or depression
iv. exposure was defined as a measure of access to green and/or blue spaces.

Outcomes
We focused our search on studies where the primary outcome was a CMD such as anxiety or depression. Our literature search focused on studies with the primary outcome defined as having or seeking care for a CMD. We removed studies that focused on severe mental health disorders as we hypothesise that exposure to the natural environment has beneficial impacts in supporting mental health and good well-being and also in the recovery of CMDs. For people who suffer with severe mental health disorders, there is no evidence to suggest that the natural environment is an effective measure

in preventing or supporting the recovery from severe mental health disorders. However, the absence of evidence is not evidence of absence.

Widely used exposures measures
Table 24 shows five exposures that we explored for this study. The primary exposures were classified as: (1) 'exposure to' and (2) 'access to' GBS.

For each exposure, we noted where we might have introduced methodologic heterogeneity.

We identified possibilities for methodological diversity within these measures. They were identified as:

1. shape of buffer
2. size of buffer
3. satellite data source
4. precise definition of exposures.

Shape of buffer

Different buffer shapes could be applied to define the immediate home environment. Euclidean, network and hexagon buffers were explored as options.

Euclidean buffers are the most widely applied in academic studies and particularly for defining exposure to 'greenness'. This is because studies tend to define exposure to 'greenness' as an ambient exposure, rather than a measure of potential access. Network buffers generally provide more accurate prediction of potential access[184] as they follow the road network. Therefore, representing the route that an individual will take to access a GBS. Hexagon buffers are applied in GIS studies to reduce computational demand. However, we decided to use circular buffers to define exposure to greenspaces in the home environment. First, circular buffers are an established method in the literature so this will enable our results to be comparable. Second, we decided our primary outcome should represent the immediate home surroundings and so our method did not need to consider the 'real' accessibility of green spaces. Finally, a circle is drawn in a GIS as a polygon with n number of sides so geometrically it is more inclusive than a hexagon polygon.

Satellite data source

Table 25 summarises possible satellite data sources that were considered for the study. Based on the spatial and temporal coverage of the data sources, we searched freely accessible online catalogues to

TABLE 24 GBS exposures considered

Exposure	Description	Studies
Exposure to green space	NDVI around the home location	Dzhambov & Dimitrova,[66] Dadvand *et al*.,[136] Sarkar *et al*.,[137] Persson *et al*.,[194] Almanza *et al*.,[278] Klompmaker *et al*.,[280]
Exposure to green space	EVI around the home location	Rugel *et al*.,[281]
Exposure to blue space	Blue space defined as coastal water	Alcock *et al*.,[39] Wheeler *et al*.,[73] Elliott *et al*.,[140] White *et al*.,[169]
Exposure to blues pace	Blue space defined as inland water	White *et al*.,[36] Gascon *et al*.,[38] Rugel *et al*.,[282] Rogerson *et al*.,[283] Grow *et al*.[284]
Access to green and blue spaces	Count of GBS from home location	Mizen *et al*.[184]

TABLE 25 Satellite data sources with spatial and temporal resolution and spectral bands available

Source	Resolution (m)	Temporal availability	Bands available
Pleiades-1A	50 cm	From 2011	Blue, green, red, NIR
SPOT 5	10	Duration of study period	Blue, green, red, NIR
SPOT 6	6	From 2012	Blue, green, red, NIR
Sentinel-1	10	From 2014	
Sentinel-2	10	From 2015	Blue, green, red, NIR
Landsat 7 & 8	30	Landsat 7 (1999) Landsat 8 (2013)	Blue, green, red, NIR
Landsat ETM	30	Duration of study period	Blue, green, red, NIR
MODIS	500	Duration of study period	Blue, green, red, NIR

scope the temporal and spatial coverage for three different satellites. Our selection criteria were images with less than 20% cloud cover during the study period.

SPOT 5
We used the ESA browser to explore the SPOT 5 (Satellite pour l'Observation de la Terre 5) data. SPOT 5 flies over any point on Earth within 26 days (i.e. monthly). A biannual mosaic will be possible with the SPOT 5 data.

Sentinel
We used the SeDAS browser to explore Sentinel data. Two national mosaics per year will be possible with the sentinel data (September to February/March to August) from 2014 to 2018.

Landsat
We used the Earth Explorer to investigate the Landsat data (https://earthexplorer.usgs.gov). Two national mosaics a year will be possible with the sentinel data (September to February/March to August) for the duration of the study period. Quarterly national coverage is not possible for the duration of the study period (2008–18) due to lack of data availability. A lack of images that have adequately low cloud cover means that it is not possible to create quarterly mosaics of Wales. However, it will be possible to create mosaics made up of images with six months of one another.

Decision
The highest-resolution satellite data was not available for the entire study period. A decision was taken to balance spatial and temporal coverage and therefore Landsat imagery was chosen as the data source to create our primary exposure.

Sources used to calculate access to green and blue spaces

Table 26 shows the sources we used to calculate access to GBS.

Precise definition of exposures

The literature is heterogenous in the precise definition of exposure to GBS. Heterogeneity arises because of methodological choices made when measuring exposure to GBS. Some studies focus on the immediate home environment, or spatially aggregate areas. Others look at the opportunity

TABLE 26 Data sources used to create GBS dataset to calculate secondary exposure (access to GBS within 1600 m)

Data source	Licence	Features
Ordnance Survey	PSMA	All natural land features extracted from the Topography Layer in OSMM
Local authority TAN	Restricted/ PSMA	Open spaces managed by local government (e.g. parks, play areas, playing fields)
NRW national inventory of woodland and trees (Lle)	OpenData	Woodland and forests
NRW national nature reserves (Lle)	OpenData	Boundaries of nature reserves
NRW forest legal boundaries (Lle)	OpenData	Boundaries of forests
NRW traditional orchards (Lle)	OpenData	Orchards

PSMA, Public Sector Mapping Agreement.

to access GBS by quantifying the proximity of GBS from the home location. There is no accepted method and when modelling environmental exposures, and the evidence base is still in the early stages of understanding methodological implications of defining exposure to GBS on health outcomes.[142,279]

Ambient exposures

NDVI has been widely applied as an indicator of exposure to greenspaces in health studies.[136,137,194,278,285] Numerous studies have calculated different buffer sizes around home locations and derived mean NDVI values for a buffer. Although NDVI is widely applied in health studies, it has been criticised as a primitive indicator of access to greenspaces. This is because this satellite-derived measure is sensitive to topographic and meteorological effects and does not contain any indicator of how to access a given greenspace. The EVI is a well-established measure in Earth observation studies, not widely applied in health studies. We decided to explore whether EVI should be used instead of NDVI as an indicator of access to greenspace. This is because EVI reduces both atmospheric and soil background noise simultaneously than NDVI. This sensitivity is an important consideration when considering the morphological context of this study.

Under objective 2, we document a sensitivity analysis between NDVI and EVI. From the results of this analysis, we decided to define the primary exposure as mean EVI within a 300 m buffer of the home location. This measure will be an indicator of exposure to greenspace in the immediate home environment.

Annual average EVI

Table 27 shows the annual average EVI aggregated to local authority level.

Table 28 shows the annual household-level mean EVI by rural/urban classification.

Opportunity to visit spaces

Generally, studies tend to look at green or blue spaces, and few have combined the two.[286] Investigating an individual's opportunity to visit a GBS in their home neighbourhood has not been as widely applied in health studies as NDVI as an indicator of exposure to greenness. In many studies, the opportunity to visit a GBS is usually defined as a Euclidean distance. A limitation of Euclidean

TABLE 27 Annual household-level mean EVI by local authority

Local authority	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Mean
Blaenau Gwent	0.25	0.27	0.29	0.22	0.27	0.32	0.21	0.34	0.38	0.22	0.28	0.36	0.29
Bridgend	0.24	0.28	0.30	0.24	0.25	0.26	0.31	0.32	0.36	0.30	0.27	0.34	0.29
Caerphilly	0.26	0.29	0.30	0.24	0.27	0.37	0.29	0.37	0.39	0.22	0.29	0.37	0.30
Cardiff	0.18	0.21	0.20	0.18	0.19	0.27	0.22	0.28	0.27	0.21	0.20	0.26	0.22
Carmarthenshire	0.26	0.33	0.34	0.33	0.26	0.35	0.40	0.40	0.44	0.38	0.33	0.41	0.35
Ceredigion	0.30	0.40	0.35	0.39	0.12	0.41	0.46	0.45	0.47	0.44	0.26	0.44	0.37
Conwy	0.21	0.29	0.26	0.28	0.03	0.28	0.31	0.38	0.38	0.36	0.30	0.31	0.28
Denbighshire	0.21	0.29	0.24	0.28	0.21	0.31	0.31	0.36	0.31	0.35	0.28	0.35	0.29
Flintshire	0.26	0.32	0.19	0.30	0.20	0.34	0.34	0.35	0.33	0.38	0.31	0.30	0.30
Gwynedd	0.27	0.36	0.30	0.35	0.03	0.39	0.34	0.45	0.42	0.42	0.38	0.40	0.34
Isle of Anglesey	0.24	0.36	0.30	0.37	−0.03	0.45	0.40	0.48	0.47	0.46	0.37	0.40	0.36
Merthyr Tydfil	0.25	0.28	0.29	0.23	0.23	0.33	0.29	0.37	0.38	0.33	0.28	0.36	0.30
Monmouthshire	0.30	0.34	0.34	0.31	0.29	0.36	0.33	0.33	0.46	0.27	0.34	0.39	0.34
Neath Port Talbot	0.22	0.25	0.29	0.21	0.20	0.30	0.29	0.33	0.35	0.30	0.28	0.34	0.28
Newport	0.21	0.23	0.25	0.21	0.24	0.29	0.26	0.28	0.32	0.25	0.23	0.29	0.25
Pembrokeshire	0.27	0.35	0.36	0.35	0.27	0.39	0.41	0.41	0.45	0.42	0.36	0.41	0.37
Powys	0.34	0.40	0.40	0.36	0.33	0.45	0.41	0.47	0.52	0.42	0.41	0.47	0.41
Rhondda Cynon Taf	0.23	0.28	0.30	0.22	0.24	0.32	0.27	0.37	0.36	0.29	0.30	0.37	0.30
Swansea	0.21	0.26	0.28	0.22	0.20	0.31	0.30	0.29	0.35	0.22	0.26	0.32	0.27
Torfaen	0.26	0.29	0.30	0.24	0.27	0.36	0.27	0.28	0.38	0.17	0.29	0.36	0.29
Vale of Glamorgan	0.23	0.28	0.26	0.24	0.24	0.27	0.30	0.28	0.33	0.26	0.22	0.30	0.27
Wrexham	0.28	0.31	0.26	0.30	0.14	0.35	0.34	0.34	0.31	0.38	0.30	0.26	0.30

TABLE 28 Annual household-level mean EVI by rural/urban classification

Rurality	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Mean
Rural	0.30	0.37	0.35	0.35	0.22	0.40	0.40	0.44	0.46	0.41	0.37	0.43	0.37
Urban	0.21	0.25	0.25	0.22	0.20	0.30	0.27	0.31	0.33	0.26	0.25	0.30	0.26

distances is that they do not account for topographic morphology and can result in environmental exposures being lost or masked.[287] For example, rivers, railway lines and motorways can impact on the accessibility of a GBS. This can lead to erroneous predictions of opportunity to visit a GBS. Network distances have been found to be more accurate representations of travel routes compared to Euclidean distances.[184]

Defining access

Defining the point of access to a GBS is another methodological decision to make. Centroids of polygons are the most widely implemented representation of an access point to a GBS. The Ordnance Survey's Greenspace dataset includes access points. However, this dataset could not be relied upon to provide accurate access points for every GBS in this national study.

We identified we could not define an access point as where the road network intersected with the boundary of a greenspace. This was because in applying our detailed typology, some GBS nested within larger spaces were assigned access points more than 1 km away from their actual location or point of access (*Figure 30*). The decision was taken to represent access to each GBS type as four points by calculating a bounding box around each space (*Figure 31*) and then calculating north, east, south and west points to represent opportunity to access a GBS.

Land cover change

To understand the most appropriate definition of opportunity to access GBS in our study, we undertook a sensitivity analysis to understand the amount of land cover change we would expect during our study.

To investigate the degree of land cover change in Wales, we compared land cover maps for 2007 and 2015. We downloaded CEH 25 m resolution land cover maps for 2007 and 2015 from Digimaps with spatial coverage for the whole of Wales. The land cover maps provide coverage for the UK and are derived from satellite images and digital cartography. The land classifications are based on UK Biodiversity Action Plan broad habitats classes. The raster dataset gives the most likely broad habitat for each pixel.

We reclassified the land classifications into binary groups: natural and urban. A raster calculation was then undertaken to determine where land use classifications had changed. Analysis was undertaken for the whole of Wales and subanalyses were undertaken for five local authorities: Swansea, Cardiff, Powys, Pembrokeshire and Wrexham. These local authorities were chosen to represent rural, semi-rural and urban areas in Wales. We define change as a change in the land-cover classification of a pixel.
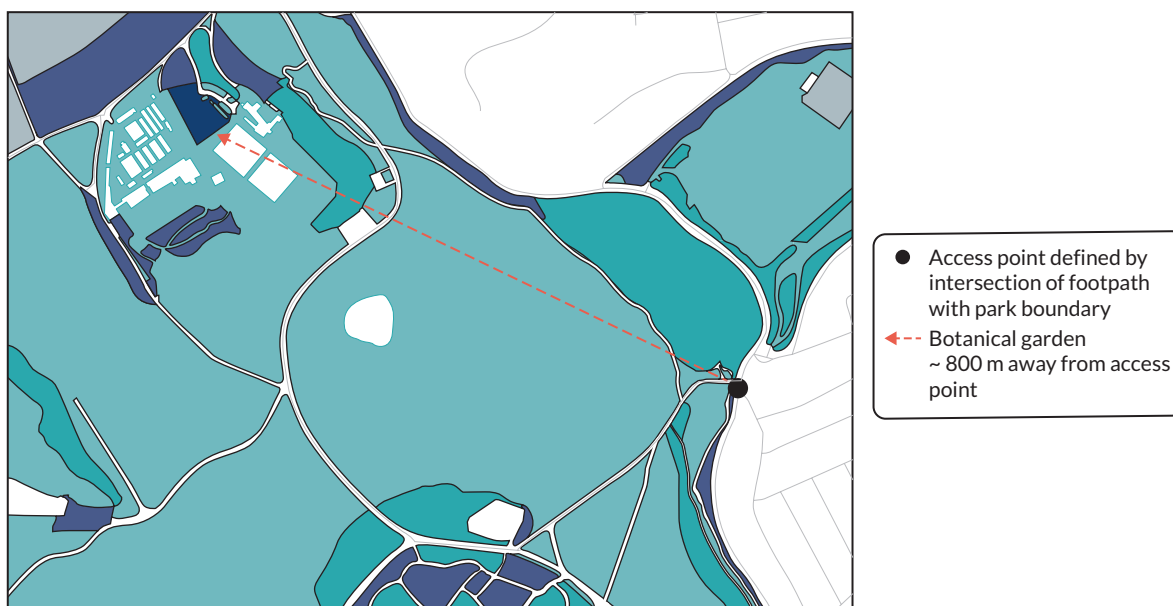

FIGURE 30 The impact of generalising access to GBS for spaces nested within other GBS spaces.
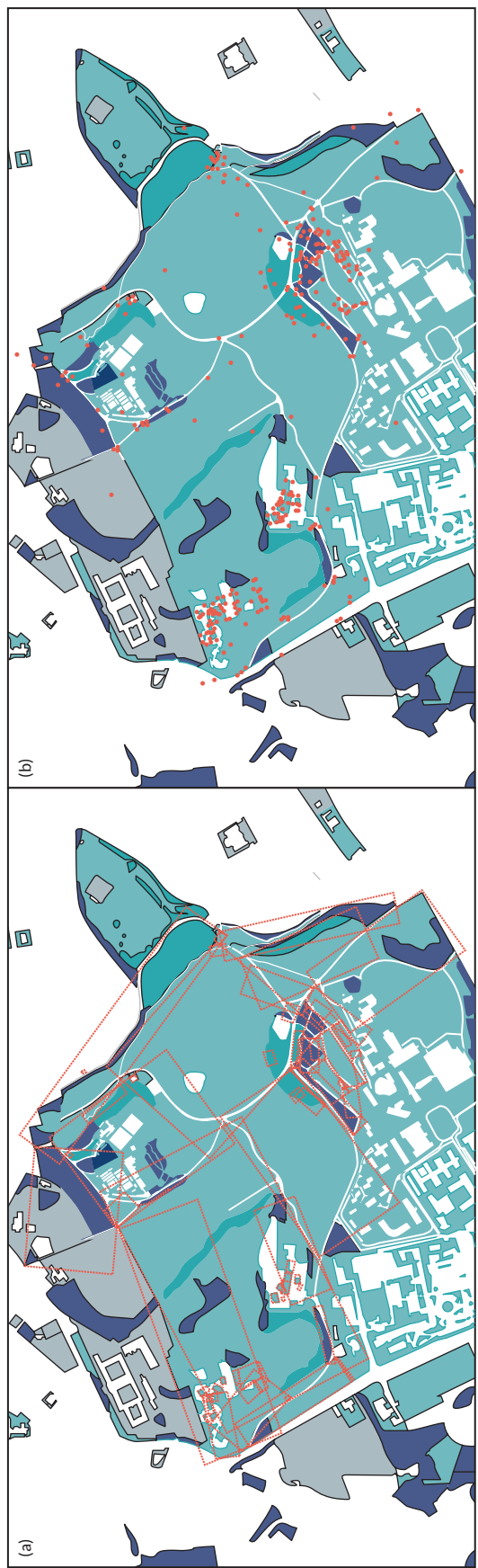
FIGURE 31 Deriving access points for GBS (a) create bounding boxes for each individual GBS (b) create north, east, south, west points based on bounding box geometry.

TABLE 29 Area of land classification changing from urban to natural, natural to urban and no change by km² and percentage change

Type of change	Area (km²)	Percentage change (%)
Urban → natural	212,554	0.81
Natural → urban	448,372	1.72
No change	25,458,852	97.47
Swansea		
Urban → natural	12,612	1.75
Natural → urban	25,996	3.60
No change	683,723	94.66
Cardiff		
Urban → natural	6427	2.20
Natural → urban	16,288	5.58
No change	269,185	92.22
Powys		
Urban → natural	33,679	0.40
Natural → urban	48,972	0.58
No change	8,296,433	99.01
Pembrokeshire		
Urban → natural	17,875	0.86
Natural → urban	45,700	2.21
No change	2,007,790	96.93
Wrexham		
Urban → natural	6414	0.50
Natural → urban	23,155	1.79
No change	1,263,521	97.71

Table 29 highlights that overall, 97% of land cover does not change in Wales during 2007–15. Urban local authorities (Swansea and Cardiff) see the largest changes; 5.44% and 7.88% of land changes, respectively. The most rural local authority sees the smallest land change and the lowest rate of urbanisation of 0.58%.

Figure 32 shows that the rate of land change from natural to urban is greater for every local authority compared with land change from urban to natural. Urban local authorities have the greatest rate of change from natural to urban land. *Figure 33* highlights that the most dominant category of change is no change for the whole of Wales and the five local authorities. *Figure 34* highlights that changes in land cover, whether urban to natural or natural to urban, are spatially heterogeneous and occur in small 'pockets' across the local authorities. Arable areas becoming grassland and grassland becoming wooded are the largest categories of change (*Figure 35*). Urban to natural land changes are relatively small.
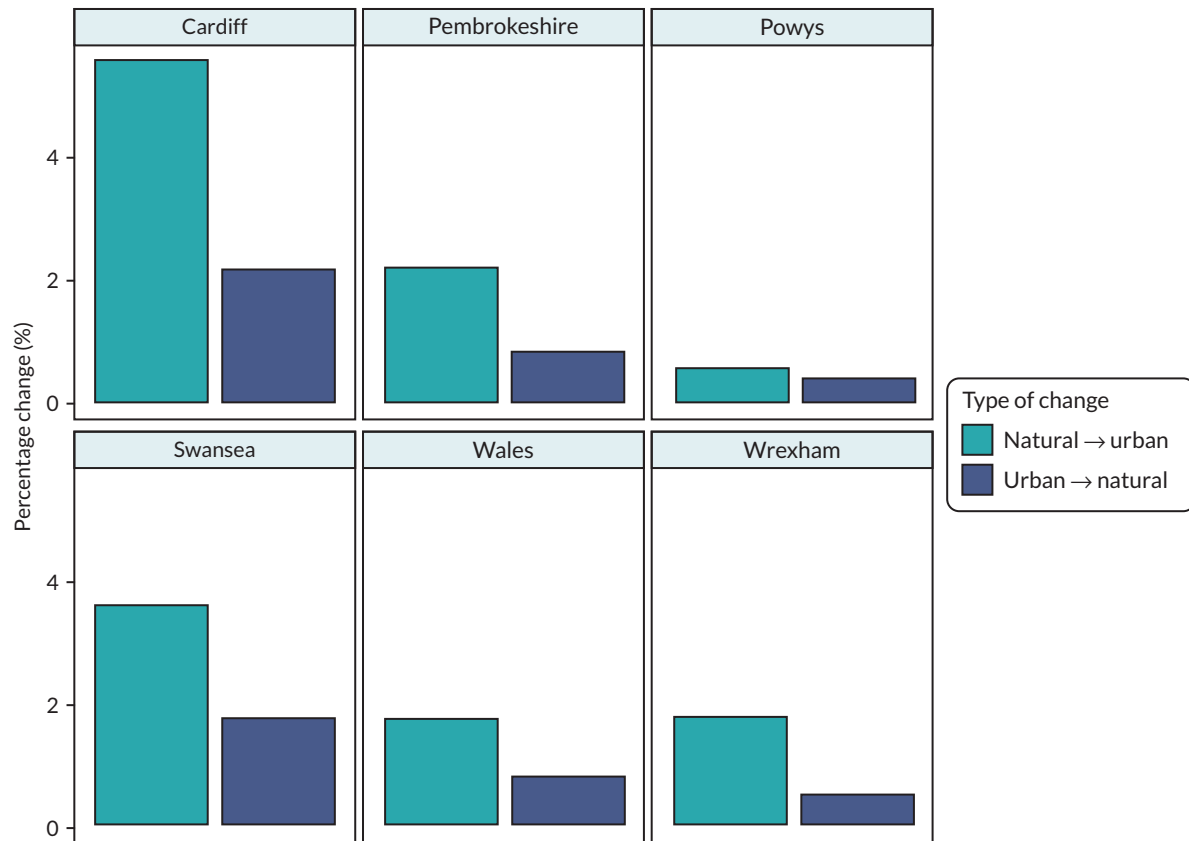
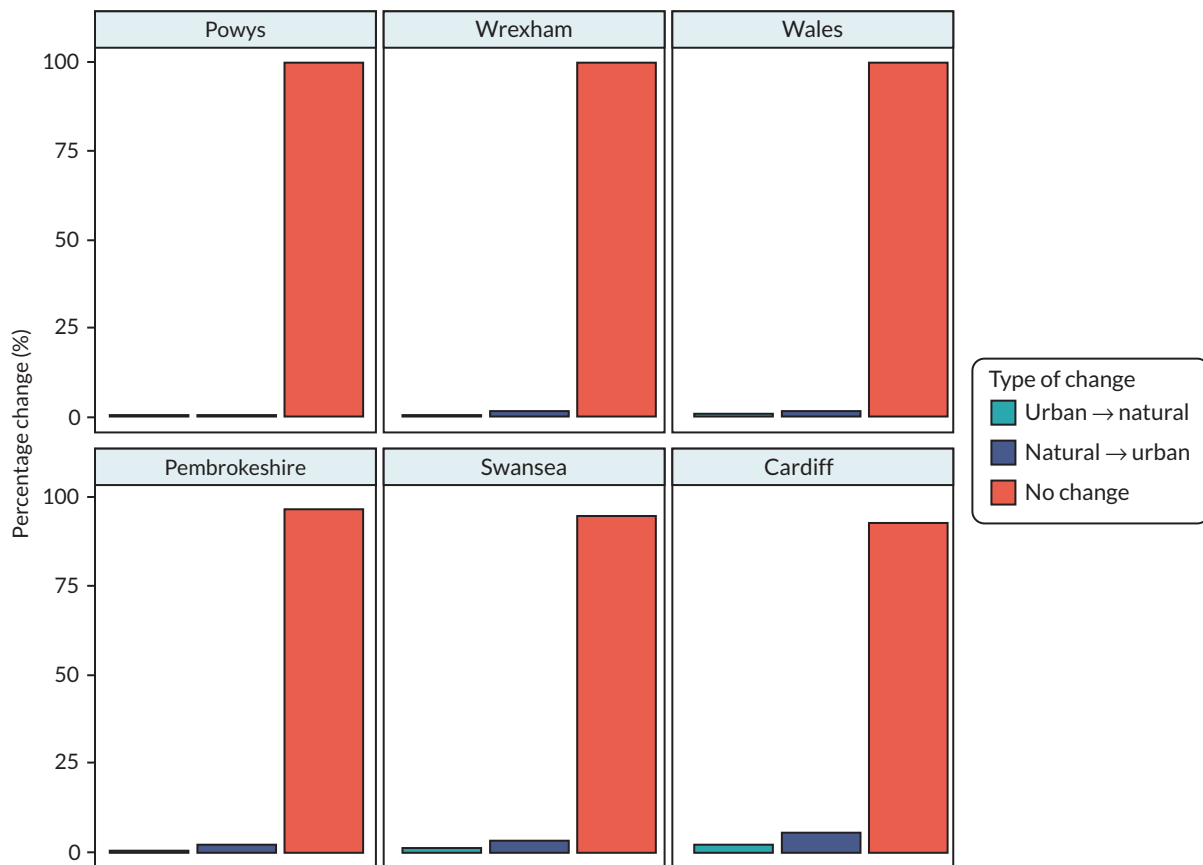**FIGURE 32** Type of land cover change by local authority.

FIGURE 33 Rate of land cover change by type and rate of no change.
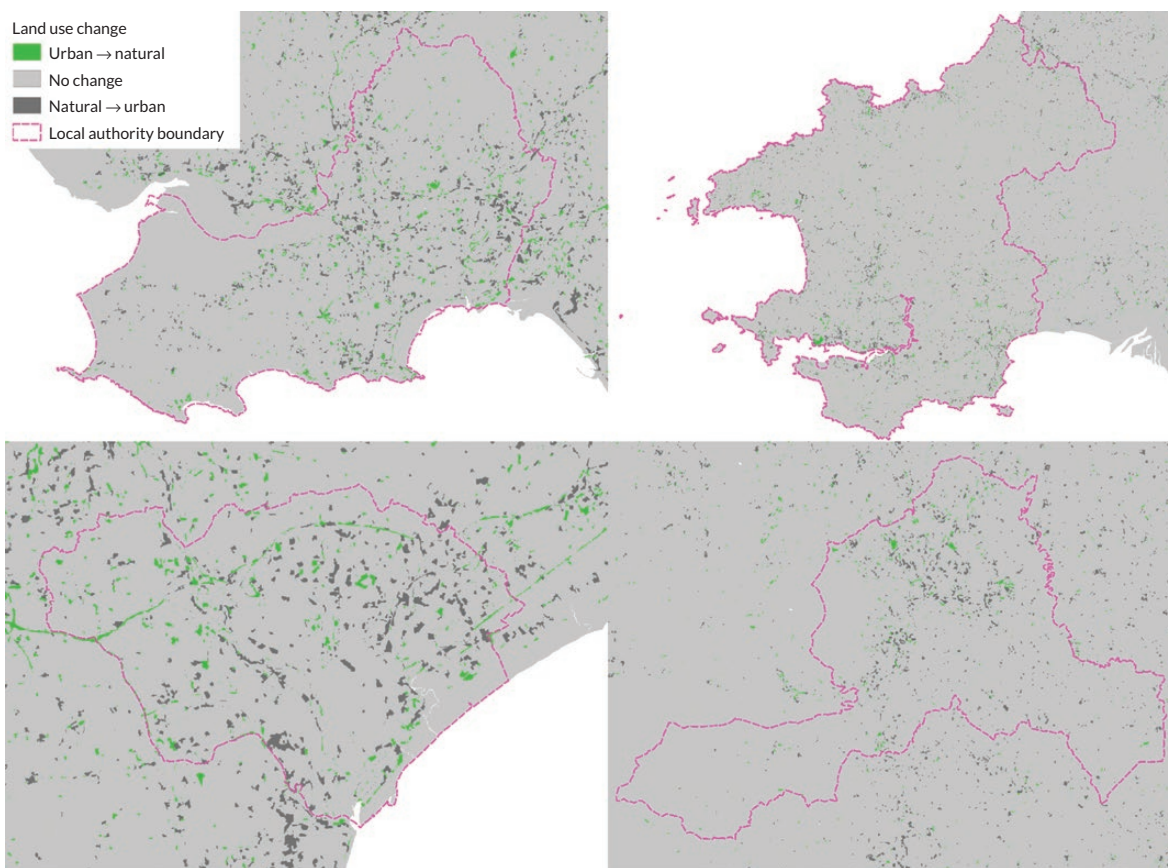
FIGURE 34 Spatial pattern of land cover change in local authorities.

The results of this land use analysis suggest that the overall amount of change in landcover in Wales is small.

Local authority data

Through previous projects,[108,288,289] the research team had extensive experience in collecting and working with administrative datasets. The research team has built up a rapport with local planning authorities in Wales and significant time was allocated to data collection.

Data were collected from local planning authorities from July 2018 to December 2018. An email request was sent in July 2018 to each local planning authorities contact in every Welsh local authority. The email gave a brief explanation of the research project and requested access to their TAN 16 data.

After one month the email was resent to the contacts who had not replied to the initial email sent in July 2018. If no response was received from this email, local planning authority general enquires were contacted by email. Of the 22 local planning authorities contacted, 14 provided their TAN 16 open space assessment. The remaining eight authorities were unable to provide any data for this request.

Recording data on managed outdoor spaces is a legal requirement for local authorities but there is no guidance on data collection or data management so local authorities provided the data in different formats, spatial resolutions, projections and with a variety of classifications for GBS. The heterogeneity in the data resulted in a challenging process of cleaning and harmonising the datasets. The open space assessments were cross-sectional and because of their limitations were used to supplement the OSMM data.
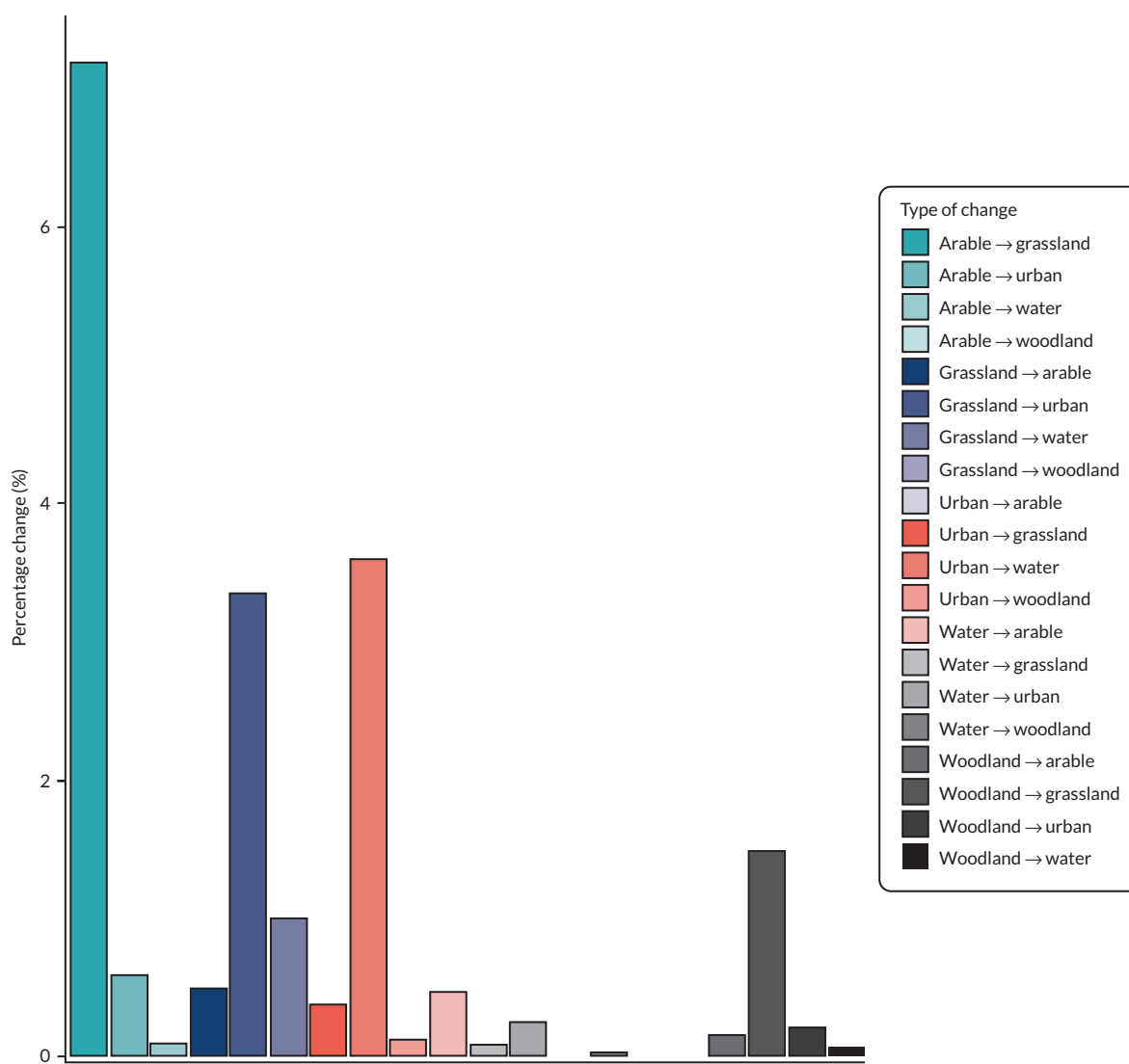
FIGURE 35 Land cover change by non-binary classifications.

Sensitivity analyses

We report and explain the sensitivity analyses we undertook to decide whether to use EVI or NDVI as the primary measure of ambient greenness exposure.

Normalised difference vegetation index

Calculating NDVI from satellite images is achieved using the following formula (*Equation 2*):

$$NDVI = (NIR - RED) / (NIR + RED) \qquad (2)$$

NDVI works using the fact that chlorophyll absorbs visible light far more readily than near-infrared (NIR) for use in photosynthesis. Therefore, a greater percentage of NIR is reflected than visible light. This means that NIR band satellite images will show areas containing chlorophyll with high values and those without having lower values. As NDVI is normalised, the range of NDVI values is between −1 and 1; however, in practice, −0.3 to 0.85 is typical.

Enhanced vegetation index

Where NDVI is chlorophyll sensitive, EVI is more responsive to canopy structural variations including leaf area index, canopy type, plant physiognomy and canopy architecture. The two vegetation indices

complement each other in global vegetation studies and improve upon the detection of vegetation changes and extraction of canopy biophysical parameters.

There are two options in the calculation of EVI. The first uses three bands in the calculation of vegetation index using the following formula:

$$EVI = G * \frac{NIR - RED}{NIR + C1 * RED - C2 * BLUE + L}$$

(3)

Where coefficients taken from the NASA MODIS-EVI algorithm: G = 2.5, L = 1, C1 = 6, C2 = 7.5.

$$EVI = 2.5 * \frac{NIR - RED}{NIR + 6 * RED - 7.5 * BLUE + 1}$$

(4)

Another option for calculating EVI uses two bands, red and NIR. This is known as two-band EVI.

$$EVI2 = 2.5 * \frac{NIR - RED}{NIR + 2.4 * RED + 1}$$

(5)

Enhanced vegetation index compared with enhanced vegetation index 2

EVI2 is advantageous as it uses only red and infrared bands but provides the benefits of EVI over NDVI as previously discussed. The calculation of values is therefore less computationally costly and when using data from satellites without the blue spectral band it can still be computed. *Figure 36* demonstrates that EVI and EVI2 are extremely similar and can be used interchangeably with no effect on the accuracy of results.
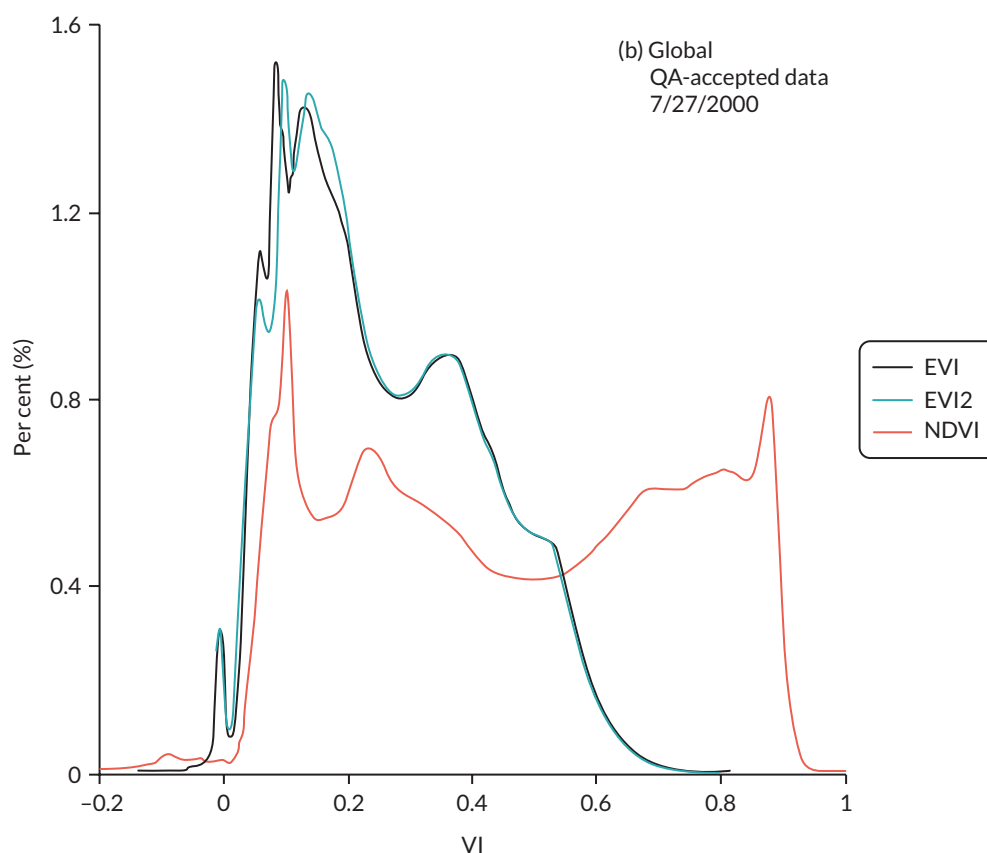
FIGURE 36 Development of a two-band EVI without a blue band. Jiang, Zhangyan & Huete, Alfredo & Didan, K & Miura, Tomoaki. (2008).

Comparison of household-level enhanced vegetation index and normalised difference vegetation index values

We calculated average NDVI and EVI values for a 300 m radius around 100,000 residential locations in South Wales (*Figures 37–40*).

Differences between normalised difference vegetation index and enhanced vegetation index

Figures 37–39 highlight that although the NDVI and EVI values produce very similar patterns, EVI values are less variable. *Figure 39* suggests that NDVI is more susceptible to seasonal variation than EVI. This is shown by the gradient of the linear best fit line which is steeper for EVI. EVI is known to produce more stable results than NDVI during sunlight, incidence angle and soil variation corrections. However, the EVI calculation is not possible for all satellites (e.g. SPOT 5 images do not include a blue band). EVI2, however, overcomes this problem – using only red and infrared bands. NDVI is a reflection of chlorophyll concentration, which is potentially easier to interpret results. Considering our study area (Wales, UK), where there is great topographical variation, EVI2 is the most appropriate measure for our study as it is less susceptible to factors that will be influenced by topographic seasonal factors.
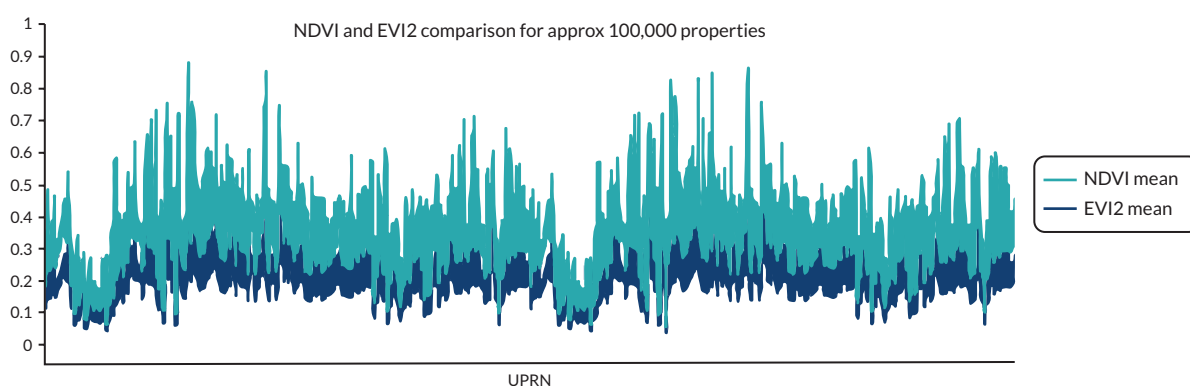
FIGURE 37 NDVI and EVI2 data points. UPRN, unique property reference number. Each individual household is represented by a UPRN.
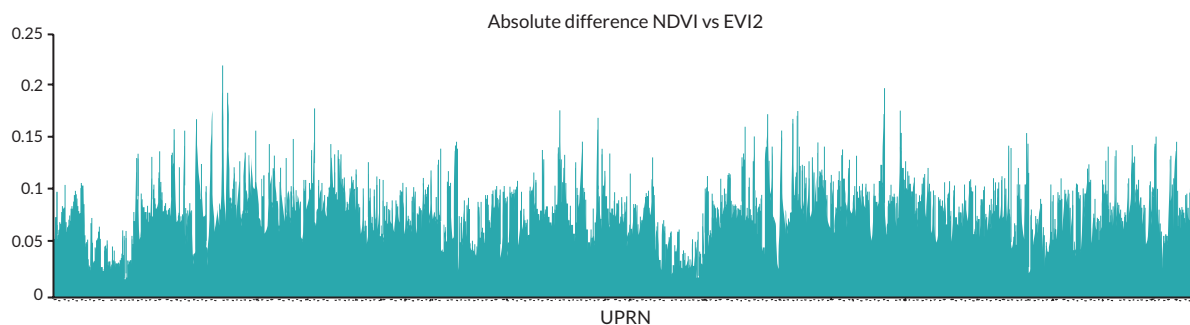
FIGURE 38 Difference between NDVI and EVI2 values. UPRN, unique property reference number. Each individual household is represented by a UPRN.
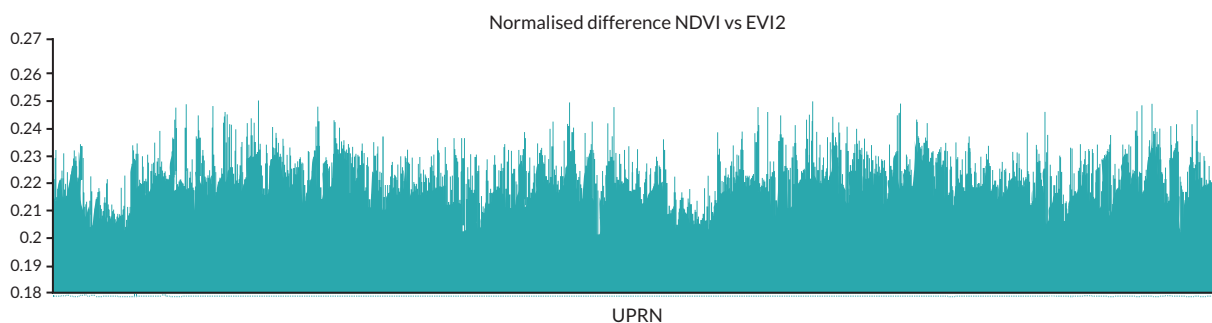
FIGURE 39 Normalised difference between NDVI and EVI2 values. UPRN, unique property reference number. Each individual household is represented by a UPRN.
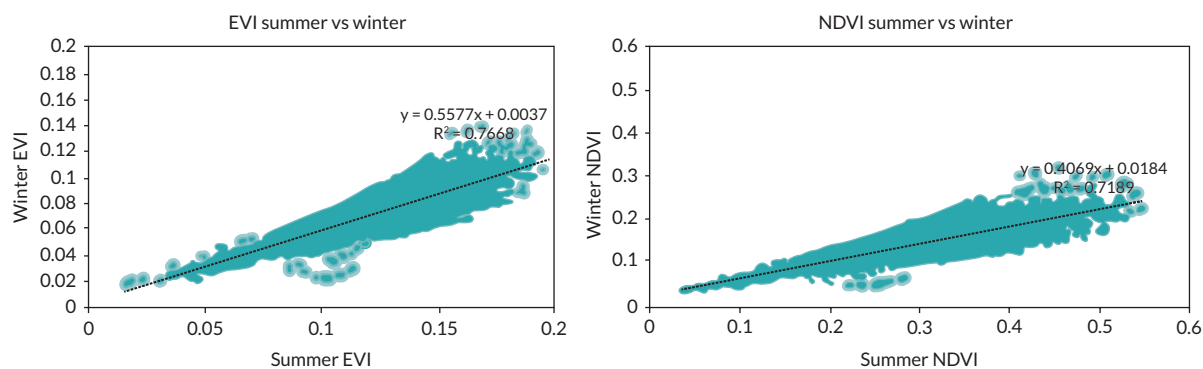
FIGURE 40 Household level seasonal differences between NDVI and EVI2 values.

Decisions

1. Satellite data source:
   a. We obtained and processed Landsat 7 and 8 data to derive EVI values for households in Wales from 2008 to 2019.
2. Shape of buffer:
   a. We will use a circular buffer instead of a hexagon or a square. This is because: (a) a circular buffer is widely applied in the literature; (b) the way a circle is drawn in a GIS is a polygon with *n* number of sides, so it is more inclusive than a hexagon.
3. Precise definition of exposures:
   - Primary exposure: annual mean EVI within 300 m of each household location
   - Secondary exposure:
     a) Annual count of GBS within 800 m of each household location
     b) Annual count of GBS within 1600 m of each household location.

Conclusion

This work has identified methodological heterogeneity in the evidence base. We undertook exploratory analysis of data sources and sensitivity analysis of methodological choices to prepare primary and secondary exposure outcomes that are scientifically robust and will also be translatable for policy and practice. This work has identified methodological heterogeneity in the evidence base. We have taken a pragmatic approach to address data availability and heterogeneity in definitions.

Resources

Our commented code which we used to generate the primary and secondary exposures is stored in https://gitlab.com/envhe/BGS.git.

Appendix 4 Sensitivity analysis

Time aggregated analysis
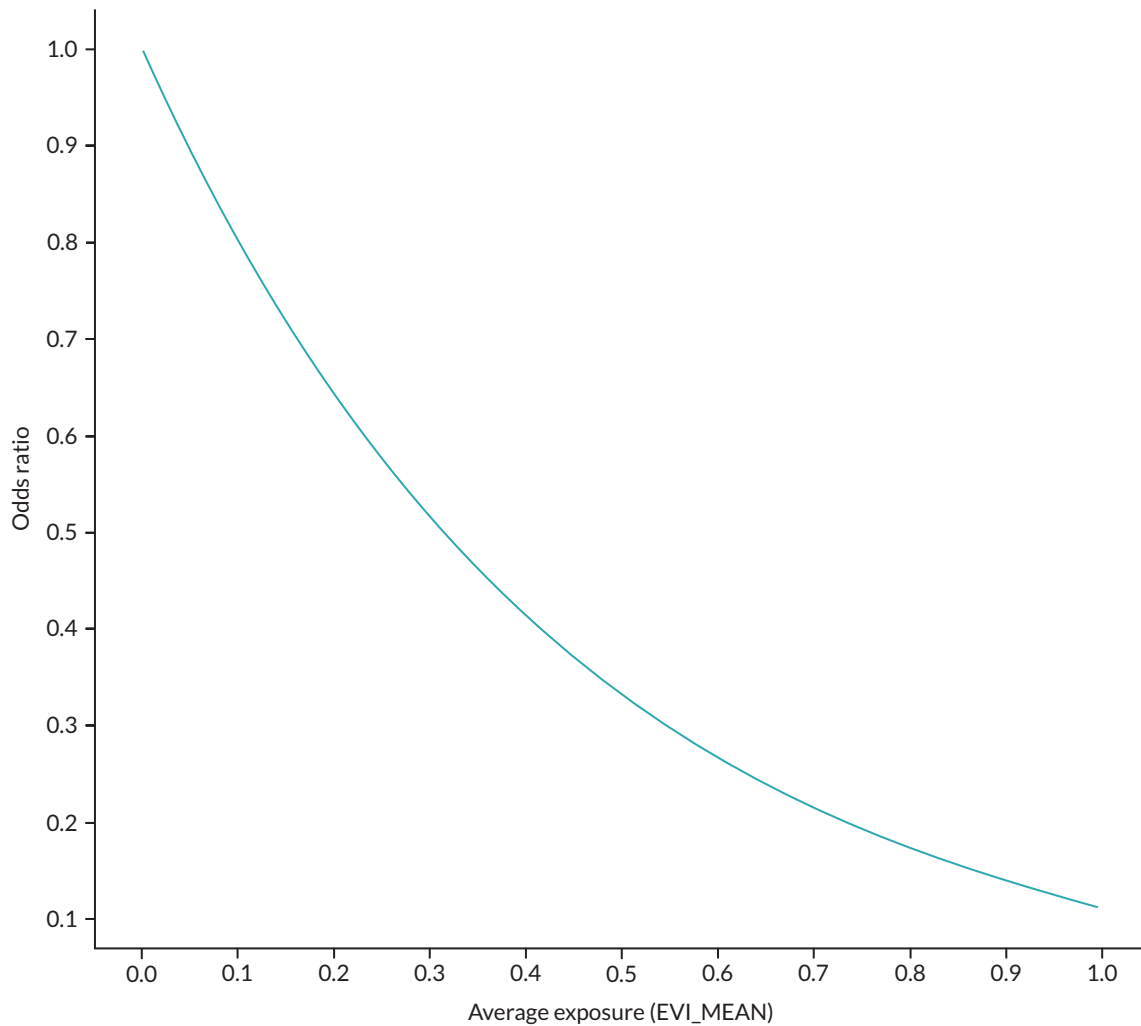
FIGURE 41 Odds ratio for CMD compared with standardised EVI.
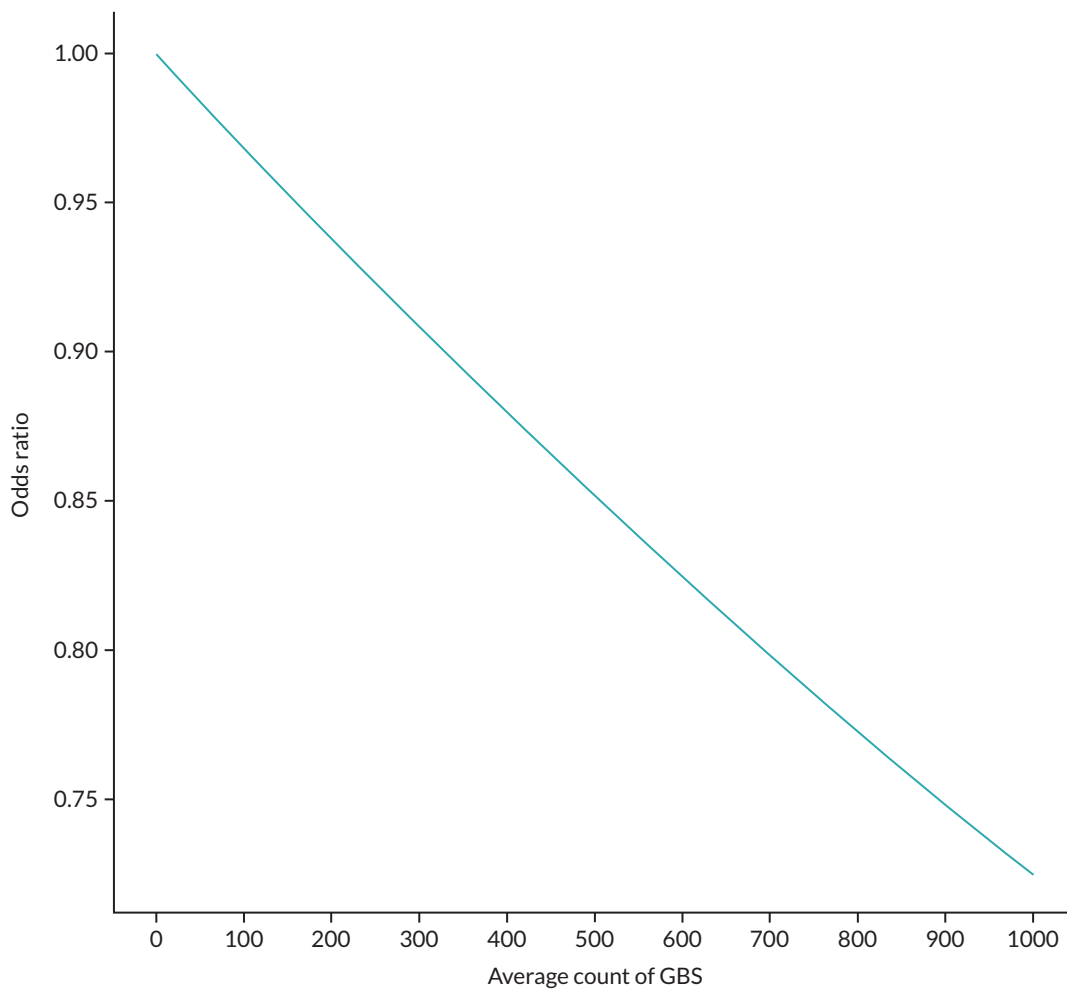
FIGURE 42 Odds ratio for CMD compared with secondary exposure, average count of GBS.

TABLE 30 Variable definition

Variables	Description	Type	Treatment	Source
CMD	Is the individual identified as having sought care for a CMD in general practice?	Primary outcome	Binary	WLGP/CMD Algorithm
Annual EVI mean	Normalised (sum(EVI_MEAN)/(number of quarters in study – number of NULL EVI quarters)	Primary exposure	Continuous	GIS dataset
Annual NDVI mean	Normalised (sum(NDVI_MEAN)/(number of quarters in study – number of NULL NDVI quarters)	Secondary exposure	Continuous	GIS dataset
Annual all counts	Normalised (sum(ALL_COUNTS)/number of quarters in study	Primary access exposure	Continuous	GIS dataset
Annual all counts	Normalised (sum(NEAREST_DISTANCE)/number of quarters in study	Secondary access exposure	Continuous	GIS dataset
Nearest GBS type	Most frequent nearest type of GBS during time interval in time-aggregated analyses	Secondary access exposure	Factor	GIS dataset
Sex	Male/female	Confounder	factor	WDSD
Settlement type	Settlement type at time of CMD or most frequent settlement type if no CMD occurs	Confounder	factor	WDSD
Deprivation quintile	Deprivation quintile at time of CMD or most frequent deprivation quintile if no CMD occurs	Confounder	factor	WDSD
Age	Age group at time of CMD or most frequent age if no CMD occurs. Age grouping (years) include: 16–21; 22–30; 31–40; 41–50; 51–60; 61–70; 71–80; 81+	Confounder	factor	WDSD
Birth or death in household	Does a birth or death occur within the linked RALF at any point during time interval in time-aggregated analyses	Confounder	Binary	WDSD/ADDE
Comorbidities	Modified Charlson Comorbidity Index derived each quarter (maximum value at any point during time interval in time-aggregated analyses)	Confounder	Continuous	
Moved home?	Does not move/moves once/moves more than once	Confounder	Factor	WDSD

ADDE, annual district death extract.

TABLE 31 Nearest type of GBS by proportion of cohort

Nearest type of GBS	All		CMD		No CMD	
	n	*%*a	*n*	*%*b	*n*	*%*b
Allotments	47,933	2.0	11,737	24.5	36,196	75.5
Beach	92,224	3.9	20,190	21.9	72,034	78.1
Botanical gardens	1581	0.1	321	20.3	1260	79.7
Canal	1601	0.1	314	19.6	1287	80.4
Coniferous	77,667	3.3	15,978	20.6	61,689	79.4
Deciduous	701,959	30.0	152,401	21.7	549,558	78.3
Estuary	365	0.0	71	19.5	294	80.5
Grassland	199,479	8.5	44,202	22.2	155,277	77.8
Harbour/dock/marina	14,411	0.6	2935	20.4	11,476	79.6
Lake	2928	0.1	655	22.4	2273	77.6
Lido	670	0.0	112	16.7	558	83.3
Marsh	13,623	0.6	2816	20.7	10,807	79.3
Meadow	530	0.0	147	27.7	383	72.3
Mixed green space	4241	0.2	750	17.7	3491	82.3
Moor heath	163,899	7.0	35,570	21.7	128,329	78.3
Multiple blue spaces	111,426	4.8	22,605	20.3	88,821	79.7
Other grounds	198,256	8.5	43,396	21.9	154,860	78.1
Park	21,920	0.9	4737	21.6	17,183	78.4
Play areas	105,347	4.5	25,566	24.3	79,781	75.7
Pond	84,978	3.6	17,867	21.0	67,111	79
Quarry	5109	0.2	1144	22.4	3965	77.6
Recreational	16,062	0.7	3443	21.4	12,619	78.6
Religious grounds	143,270	6.1	30,849	21.5	112,421	78.5
Reservoir	4209	0.2	904	21.5	3305	78.5
River	129,236	5.5	28,241	21.9	100,995	78.1
School grounds	160,661	6.9	37,013	23.0	123,648	77
Sports pitches	37,547	1.6	9191	24.5	28,356	75.5
Undefined	459	0.0	84	18.3	375	81.7

a Percentage of the whole cohort in this category.
b Percentage of individuals in this category with/without a CMD.

TABLE 32 Unadjusted and adjusted logistic regression models for CMD: comparison between ambient greenness (EVI) and ambient greenness (NDVI)

Variable	(N)	(%)	Ambient greenness (EVI) model (unadjusted)			Ambient greenness (EVI) model (adjusted)			Model: NDVI (adjusted)		
			(OR)	95% CI		AOR	95% CI		AOR	95% CI	
CMD	513,239	21.9									
Variable											
EVI or NDVI annual mean		15.5	0.77	0.75	0.77	0.80	0.80	0.81	0.63	0.63	0.63
Age group (years)											
16–21 (ref)	363,314										
22–30	385,643	16.50	1.53	1.51	1.55	1.65	1.63	1.67	1.46	1.45	1.48
31–40	327,874	14.00	1.54	1.52	1.55	1.71	1.69	1.73	1.55	1.53	1.57
41–50	338,303	14.40	1.50	1.49	1.52	1.59	1.57	1.60	1.47	1.46	1.49
51–60	302,774	12.90	1.21	1.19	1.22	1.29	1.27	1.31	1.21	1.20	1.23
61–70	284,096	12.10	0.83	0.82	0.84	0.95	0.94	0.96	0.90	0.89	0.91
71–80	201,279	8.60	0.75	0.74	0.76	0.94	0.93	0.96	0.88	0.86	0.89
80+	138,308	5.90	0.64	0.63	0.66	0.90	0.88	0.92	0.80	0.78	0.81
Sex											
Male (ref)	1,193,240	51.00									
Female	1,148,351	49.00	1.76	1.75	1.77	1.84	1.83	1.85	1.83	1.82	1.84
Birth in household											
No (ref)	2,076,801	88.70									
Yes	264,790	11.30	0.88	0.87	0.89	0.75	0.74	0.75	0.76	0.75	0.77

Table 32 Unadjusted and adjusted logistic regression models for CMD: comparison between ambient greenness (EVI) and ambient greenness (NDVI) *(continued)*

	(N)	(%)	Ambient greenness (EVI) model (unadjusted)			Ambient greenness (EVI) model (adjusted)			Model: NDVI (adjusted)		
			(OR)	95% CI	95% CI	AOR	95% CI	95% CI	AOR	95% CI	95% CI
Death in household											
No (ref)	2,151,488	91.90									
Yes	190,103	8.10	0.49	0.48	0.49	0.65	0.64	0.65	0.66	0.65	0.66
Charlson Comorbidity Index			0.94	0.94	0.94	0.96	0.96	0.96	0.96	0.96	0.96
WIMD											
1 – Most deprived	473,404	20.20	1.72	1.70	1.73	1.64	1.62	1.66	1.55	1.54	1.57
2	462,492	19.80	1.40	1.38	1.41	1.38	1.37	1.40	1.33	1.32	1.34
3	480,278	20.50	1.15	1.14	1.17	1.20	1.19	1.21	1.19	1.17	1.20
4	445,002	19.00	1.03	1.02	1.04	1.08	1.06	1.09	1.05	1.04	1.06
5 – Least deprived (ref)	480,415	20.50	1.00	-	-	1.00	-	-	1.00	-	-
Moved home											
Did not move (ref)	1,095,470	61.50									
Moved once	420,815	16.30	0.75	0.75	0.76	0.70	0.69	0.70	0.70	0.70	0.71
Moved more than once	312,067	22.20	0.71	0.71	0.72	0.60	0.59	0.60	0.60	0.59	0.61

TABLE 33 Model progression – unadjusted and adjusted logistic regression model results for primary ambient exposure (EVI mean)

Variable	(N)	(%)	Model: IM_1E (unadjusted) OR	95% CI		Model: IM_2E AOR	95% CI		Model: IM_3E AOR	95% CI		Model: IM_4E AOR	95% CI		Model: IM_5E (includes settlement type)	95% CI	
CMD	513,239	21.9															
Variable																	
Normalised annual ambient exposure (EVI mean)			0.77	0.75	0.77	0.78	0.78	0.78	0.81	0.80	0.81	0.80	0.80	0.81	0.75	0.75	0.77
Age group (years)																	
16–21 (ref)	363,314	15.5															
22–30	385,643	16.5	1.53	1.51	1.55	1.52	1.50	1.54	1.53	1.51	1.55	1.65	1.63	1.67	1.64	1.62	1.65
31–40	327,874	14.0	1.54	1.52	1.55	1.64	1.62	1.65	1.64	1.62	1.66	1.71	1.69	1.73	1.69	1.67	1.71
41–50	338,303	14.4	1.50	1.49	1.52	1.60	1.58	1.62	1.62	1.60	1.63	1.59	1.57	1.60	1.57	1.55	1.58
51–60	302,774	12.9	1.21	1.19	1.22	1.34	1.33	1.36	1.36	1.34	1.38	1.29	1.27	1.31	1.27	1.26	1.29
61–70	284,096	12.1	0.83	0.82	0.84	1.01	1.00	1.02	1.03	1.02	1.04	0.95	0.94	0.96	0.94	0.92	0.95
71–80	201,279	8.6	0.75	0.74	0.76	1.03	1.01	1.04	1.05	1.03	1.06	0.94	0.93	0.96	0.93	0.91	0.94
80+	138,308	5.9	0.64	0.63	0.66	0.95	0.94	0.97	0.99	0.97	1.00	0.90	0.88	0.92	0.88	0.86	0.90
Sex																	
Male (ref)	1,193,240	51.0															
Female	1,148,351	49.0	1.76	1.75	1.77	1.80	1.79	1.82	1.82	1.81	1.83	1.84	1.83	1.85	1.84	1.83	1.85
Birth in household																	
No (ref)	2,076,801	88.7															
Yes	264,790	11.3	0.88	0.87	0.89	0.68	0.67	0.69	0.66	0.66	0.67	0.75	0.74	0.75	0.75	0.74	0.76

continued

Table 33 Model progression – unadjusted and adjusted logistic regression model results for primary ambient exposure (EVI mean) (continued)

	(N)	(%)	Model: IM_1E (unadjusted) OR	95% CI		Model: IM_2E AOR	95% CI		Model: IM_3E AOR	95% CI		Model: IM_4E AOR	95% CI		Model: IM_5E (includes settlement type) AOR	95% CI	
Death in household																	
No (ref)	2,151,488	91.9															
Yes	190,103	8.1	0.49	0.48	0.49	0.62	0.61	0.62	0.60	0.59	0.61	0.65	0.64	0.65	0.64	0.64	0.65
Charlson Comorbidity Index			0.94	0.94	0.94	0.96	0.96	0.96	0.96	0.95	0.96	0.96	0.96	0.96	0.96	0.96	0.96
WIMD																	
1 – Most deprived	473,404	20.2	1.72	1.70	1.73				1.64	1.63	1.66	1.64	1.62	1.66	1.67	1.65	1.68
2	462,492	19.8	1.40	1.38	1.41				1.38	1.37	1.40	1.38	1.37	1.40	1.37	1.36	1.39
3	480,278	20.5	1.15	1.14	1.17				1.19	1.18	1.21	1.20	1.19	1.21	1.17	1.16	1.18
4	445,002	19.0	1.03	1.02	1.04				1.07	1.05	1.08	1.08	1.06	1.09	1.04	1.03	1.05
5 – Least deprived (ref)	480,415	20.5															
Migration																	
Does not move (ref)	1,095,470	61.5															
Moves once	420,815	16.3	0.75	0.75	0.76							0.70	0.69	0.70	0.60	0.59	0.60
Moves more than once	312,067	22.2	0.71	0.71	0.72							0.60	0.59	0.60	0.70	0.69	0.70
Rurality																	
Urban	1,611,516	68.8	1.01	1.00	1.01										0.80	0.79	0.81
Rural	354,961	15.2	0.73	0.73	0.74										1.03	1.02	1.05
Town and fringe (ref)	375,114	16.0															

Models IM_1E are all univariate models; model IM_2E introduces additional individual level covariates (age group, sex, death in household, birth in household, comorbidity index); model_3E introduces deprivation; model_4E (main model used in analyses) includes migration; model_5E includes settlement type.

TABLE 34 Unadjusted and adjusted logistic regression model results for primary access measure to GBS (all counts)

Variable	N	%	Model: IM_1A (unadjusted) OR	95% CI		Model: IM_2A AOR	95% CI		Model: IM_3A AOR	95% CI		Model: IM_4A AOR	95% CI		Model: IM_5A AOR	95% CI	
CMD	513,239	21.9															
Variable																	
Normalised annualised access to GBS (all counts)			0.74	0.72	0.75	0.58	0.57	0.59	0.50	0.49	0.52	0.47	0.46	0.48	0.34	0.33	0.35
Age group (years)																	
16–21 (ref)	363,314	15.5															
22–30	385,643	16.5				1.66	1.64	1.68	1.66	1.64	1.68	1.81	1.79	1.83	1.81	1.79	1.83
31–40	327,874	14.0				1.70	1.68	1.72	1.70	1.68	1.72	1.79	1.76	1.81	1.78	1.76	1.80
41–50	338,303	14.4				1.58	1.56	1.60	1.59	1.58	1.61	1.57	1.56	1.59	1.57	1.56	1.59
51–60	302,774	12.9				1.30	1.28	1.31	1.32	1.30	1.33	1.26	1.24	1.27	1.26	1.25	1.28
61–70	284,096	12.1				0.95	0.94	0.97	0.98	0.97	1.00	0.91	0.90	0.92	0.92	0.91	0.93
71–80	201,279	8.6				0.99	0.98	1.00	1.02	1.01	1.04	0.93	0.91	0.94	0.93	0.92	0.95
80+	138,308	5.9				0.96	0.95	0.98	1.00	0.98	1.02	0.92	0.90	0.93	0.92	0.90	0.94
Sex																	
Male (ref)	1,193,240	51.0															
Female	1,148,351	49.0				1.81	1.80	1.82	1.83	1.82	1.84	1.85	1.83	1.86	1.85	1.83	1.86
Birth in household																	
No (ref)	2,076,801	88.7															
Yes	264,790	11.3				0.66	0.66	0.67	0.64	0.63	0.64	0.72	0.71	0.73	0.72	0.71	0.73
Death in household																	
No (ref)	2,151,488	91.9															
Yes	190,103	8.1				0.60	0.60	0.61	0.59	0.58	0.59	0.63	0.62	0.63	0.63	0.62	0.64
Comorbidity index						0.96	0.96	0.96	0.96	0.95	0.96	0.96	0.96	0.96	0.96	0.96	0.96

Table 34 Unadjusted and adjusted logistic regression model results for primary access measure to GBS (all counts) *(continued)*

	N	%	Model: IM_1A (unadjusted) OR	95% CI		Model: IM_2A AOR	95% CI		Model: IM_3A AOR	95% CI		Model: IM_4A AOR	95% CI		Model: IM_5A AOR	95% CI	
WIMD																	
1 – Most deprived	473,404	20.2							1.81	1.79	1.83	1.81	1.79	1.83	1.75	1.73	1.77
2	462,492	19.8							1.44	1.43	1.46	1.44	1.43	1.46	1.42	1.41	1.44
3	480,278	20.5							1.16	1.15	1.18	1.17	1.16	1.18	1.21	1.20	1.23
4	445,002	19.0							1.03	1.02	1.04	1.03	1.02	1.04	1.09	1.07	1.10
5 – Least deprived (ref)	480,415	20.5															
Migration																	
Does not move (ref)	1,095,470	61.5															
Moves once	420,815	16.3										0.69	0.69	0.70	0.69	0.69	0.70
Moves more than once	312,067	22.2										0.60	0.59	0.60	0.59	0.58	0.60
Rurality																	
Urban	1,611,516	68.8													1.09	1.08	1.10
Rural	354,961	15.2													0.78	0.77	0.79
Town and fringe (ref)	375,114	16.0															

Models IM_1A are all univariate models; Model IM_2A introduces additional individual level covariates (Age Group, Gender, Death in Household, Birth in Household, Comorbidity Index; Model_3A introduces deprivation; Model_4A (main model used in analyses) includes migration; Model_5A includes settlement type.

Appendix 5　Secondary environmental exposure analyses

Change in GBS access

TABLE 35 CMD descriptive statistics by sociodemographic characteristics at baseline

	Number of quarters per distinct CMD period per individual			Total number of quarters of CMD per individual			Number of distinct CMD periods per individual		
	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Sex									
Male	12.84	6.00	13.53	16.74	10.00	15.00	1.45	1.00	0.76
Female	14.80	8.00	14.59	19.93	14.00	15.94	1.56	1.00	0.86
Age group (years)									
16–21	7.55	5.00	7.01	11.00	7.00	9.87	1.48	1.00	0.78
22–30	11.20	6.40	11.34	17.05	12.00	14.24	1.67	1.00	0.92
31–40	14.76	8.00	14.22	20.64	16.00	15.94	1.62	1.00	0.88
41–50	17.14	9.50	15.81	22.45	18.00	16.7	1.55	1.00	0.84
51–60	19.18	11.00	16.90	23.61	20.00	17.24	1.45	1.00	0.78
61–70	18.63	10.50	16.82	22.55	18.00	17.21	1.41	1.00	0.76
71–80	15.99	9.00	15.15	19.03	13.00	15.84	1.34	1.00	0.68
81+	12.92	8.00	12.22	14.66	10.00	12.97	1.20	1.00	0.5
WIMD									
1 – Most deprived	13.55	7.00	13.97	17.90	11.00	15.48	1.49	1.00	0.80
2	13.56	7.00	13.96	17.92	12.00	15.44	1.50	1.00	0.81
3	14.02	7.00	14.26	18.54	12.00	15.66	1.51	1.00	0.82
4	14.23	7.00	14.31	19.00	13.00	15.75	1.53	1.00	0.83
5 – Least deprived (ref)	14.81	8.00	14.61	19.99	14.00	15.93	1.57	1.00	0.86

TABLE 36 Change in mean EVI and the number of potentially accessible GBS upon moving home

| | Percentage of all house moves | |
	Mean EVI	Number of potentially accessible GBS
No change	0.2	7.2
Increase	50.9	46.5
Decrease	48.9	46.3

House moves where both residences had EVI/number of GBS metrics available included.

Appendix 6 Exposures and covariates

Summary of metrics

TABLE 37 Summary of green/blue space exposure metrics

Type of exposure	Definition	Measure & source	Description
Use (visits)	Intentionally visiting or spending time in GBS for active or passive activities. We focus on leisure time use of GBS, as opposed to professional (e.g. farming) or incidental (e.g. while travelling)	Estimated weekly outdoor leisure time (NSW)	Calculated by dividing number of visits to the Welsh outdoors in the last four weeks by four, and multiplying by time spent doing the 'main activity' to the outdoors during the most recent visit. Capped at 420 minutes per week in analysis (see text)
Access/ proximity	Measures to represent people's potential to visit GBS in their locality within a short walk	Distance to nearest potentially accessible GBS (SAIL linkage)	Distance from residential address to nearest access point for potentially accessible GBS within 1600 m of the home (see *Chapter 3*)
Ambient green-ness	The presence of green space and other green infrastructure (e.g. street trees, domestic gardens) without assumption of availability for public use	EVI within 300 m buffer around home (SAIL linkage)	Measure of vegetation within 300 m of the home (yearly mean); derived from remotely sensed measurements (see *Chapter 3*)

Directed acyclic graphs
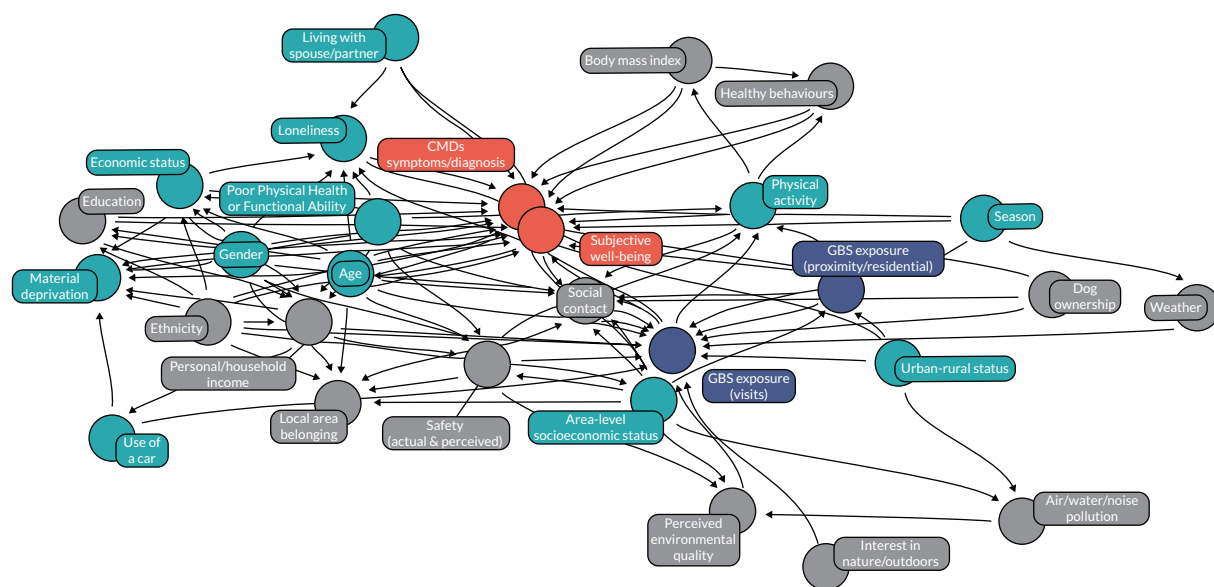
FIGURE 43 Full directed acyclic graph. **Orange**: outcomes; **green**: exposures; **blue**: potential confounders/mediators (data available); **grey**: potential confounders/mediators (data not available or availability limited).
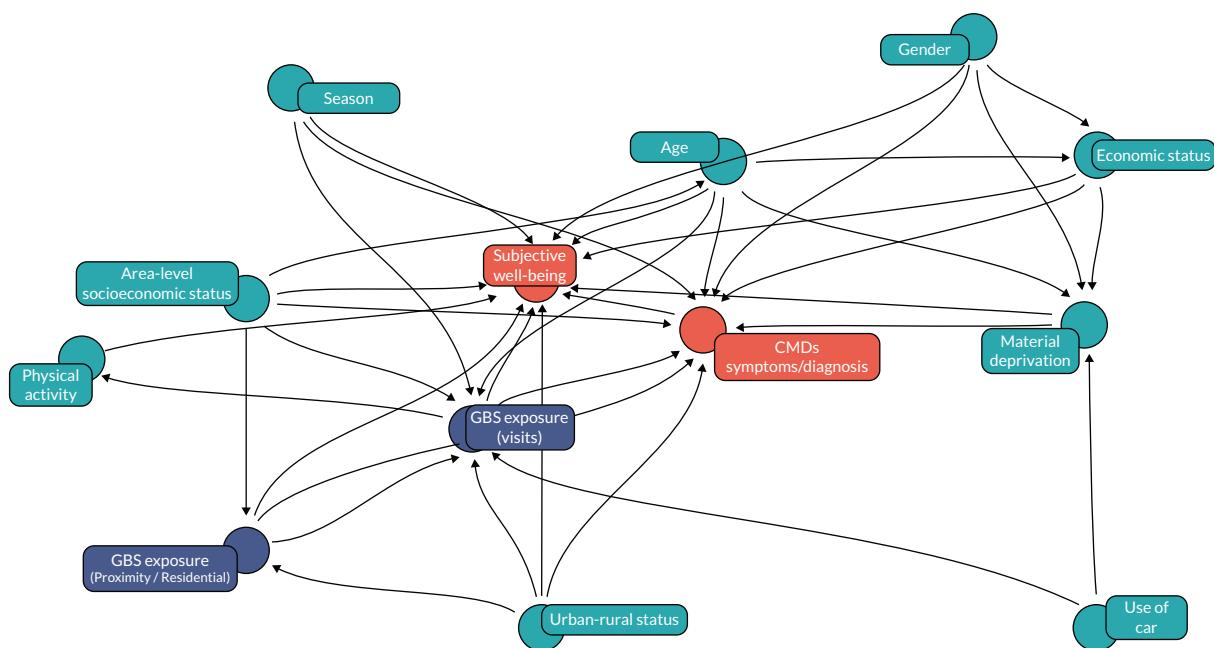
FIGURE 44 Directed acyclic graph showing available variables only (extract from *Figure 43*). **Orange**: outcomes; **green**: exposures; **blue**: potential confounders/mediators.

TABLE 38 Sample size details at each stage of the data linking process

	Missing (*n*)	Sample size (*n*)
NRW subsample of the NSW for 2016/17 and 2018/19 with ALFs (individual linking field) present		11,378
RALF (residential linking field)	634	10,744
CMD	2441	8303
WIMD	36	8267
EVI	298	7969
Time outdoors	68	7901
SWB – WEMWBS and life satisfaction	WEMWBS (*n* = 1299) Life satisfaction (*n* = 21) Unique rows (*n* = 1310)	6591
Covariates	Material deprivation (*n* = 56) Economic status (*n* = 12) Sex (*n* = 3) Unique rows (*n* = 70)	6521
Proximity to nearest GBS (m)	550	5971

Appendix 7 Analyses underpinning modelling decisions

Table 39 summarises the approach used to model predictors based on preliminary analyses described in the main text of *Chapter 7*. This is followed by details of the results from these analyses.

Generalised additive models

Generalised additive models (with thin plate regression splines; 'ts' smooth function) were the first step in the analysis process. These were used to explore the shape of the relationship between the key predictors and the outcomes. Analyses were carried out in RStudio (version 1.4.1103) with R (version 4.0.54) and packages 'mgcv' (Wood 2021, version 1.8-34) and 'gratia' (version 0.6.0). An effective degrees of freedom (edf) of 1 would indicate a linear relationship, while an edf of 2 indicates a quadratic relationship.

Enhanced vegetation index predicting Warwick–Edinburgh Mental Well-being Scale

EVI was significantly related to WEMWBS with a U-shaped relationship function (edf = 2.351, p = 0.026) indicating an approximately quadratic relationship (*Figure 45*).

The GAM indicated no significant relationship between EVI and WEMWBS for those in material deprivation (edf = 0.01, p = 0.413) and a significant relationship between EVI and WEMWBS for those not in material deprivation (edf = 2.50, p = 0.018; *Figure 46*).

For those in urban areas, the GAM indicated a significant approximately linear negative relationship between EVI and WEMWBS (edf = 0.85, p = 0.013), while for those in village, hamlet and isolated dwellings it was approximately a quadratic relationship (edf = 1.84, p = 0.082) and for those in towns and urban fringes there was no relationship (edf = 0.01, p = 0.840; *Figure 47*).

Generalised linear modelling decision
EVI predicting WEMWBS was modelled with both a linear and a quadratic term in subsequent GLMs.

Enhanced vegetation index predicting life satisfaction

A similar approximately quadratic relationship was found for EVI predicting life satisfaction for the full sample (edf = 2.12, p = 0.029; *Figure 48*). Relationships were approximately quadratic for those not in material deprivation (edf = 2.07, p = 0.062) and nearer linear for those in material deprivation (edf = 0.80, p = 0.027; *Figure 49*). Although there were no significant relationships by urban status (Urban: edf < 0.01, p = 0.412; town and fringe: edf < 0.01, p = 0.372; Village, hamlet and isolated dwellings: edf = 0.1, p = 0.511).

Generalised linear modelling decision
EVI predicting life satisfaction was modelled with both a linear and a quadratic term in subsequent GLMs.

TABLE 39 Summary of the approach used to model the predictors by exposure type and outcome in further GLM analyses

Predictor	Outcome	Modelling terms
EVI	WEMWBS	Linear and quadratic, continuous
EVI	Life satisfaction	Linear and quadratic, continuous
EVI	CMD status	Categories; < 0.2, 0.2 – 0.4 and > 0.4
Proximity to nearest GBS	WEMWBS	Categories; 0 – < 100, 100 - < 300, 300 – < 500 and > 500 m
Proximity to nearest GBS	Life satisfaction	Categories; 0 – < 100, 100 - < 300, 300 – < 500 and > 500 m
Proximity to nearest GBS	CMD status	Linear and quadratic, continuous
Time outdoors	WEMWBS	Linear and quadratic, continuous
Time outdoors	Life satisfaction	Linear, continuous
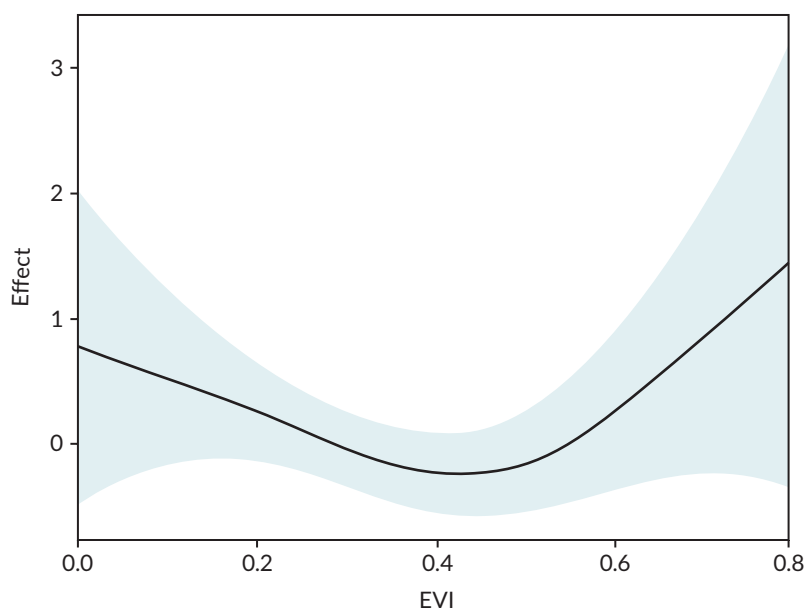
FIGURE 45 The smoothed function of EVI on WEMWBS. Model includes WIMD, sex, age, economic status, material deprivation, car use, season and wave.
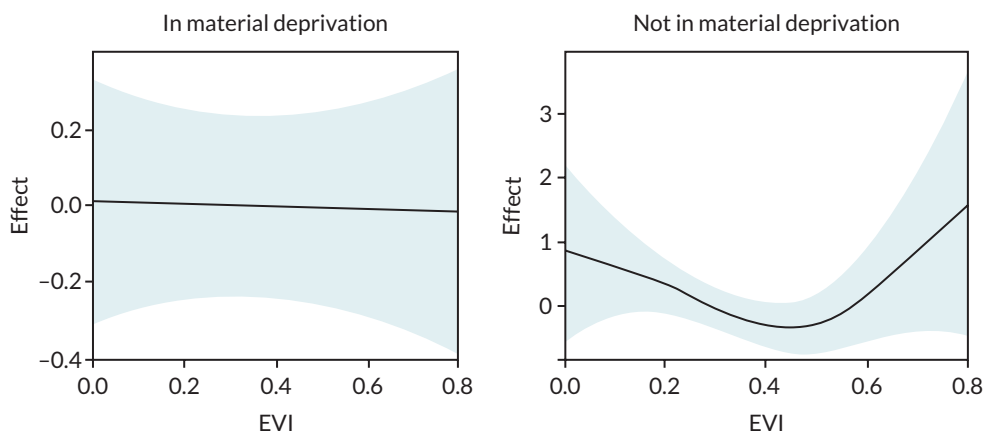
FIGURE 46 The smoothed function of EVI on WEMWBS by deprivation. Model includes WIMD, urban status, sex, age, economic status, material deprivation, car use, season and wave.
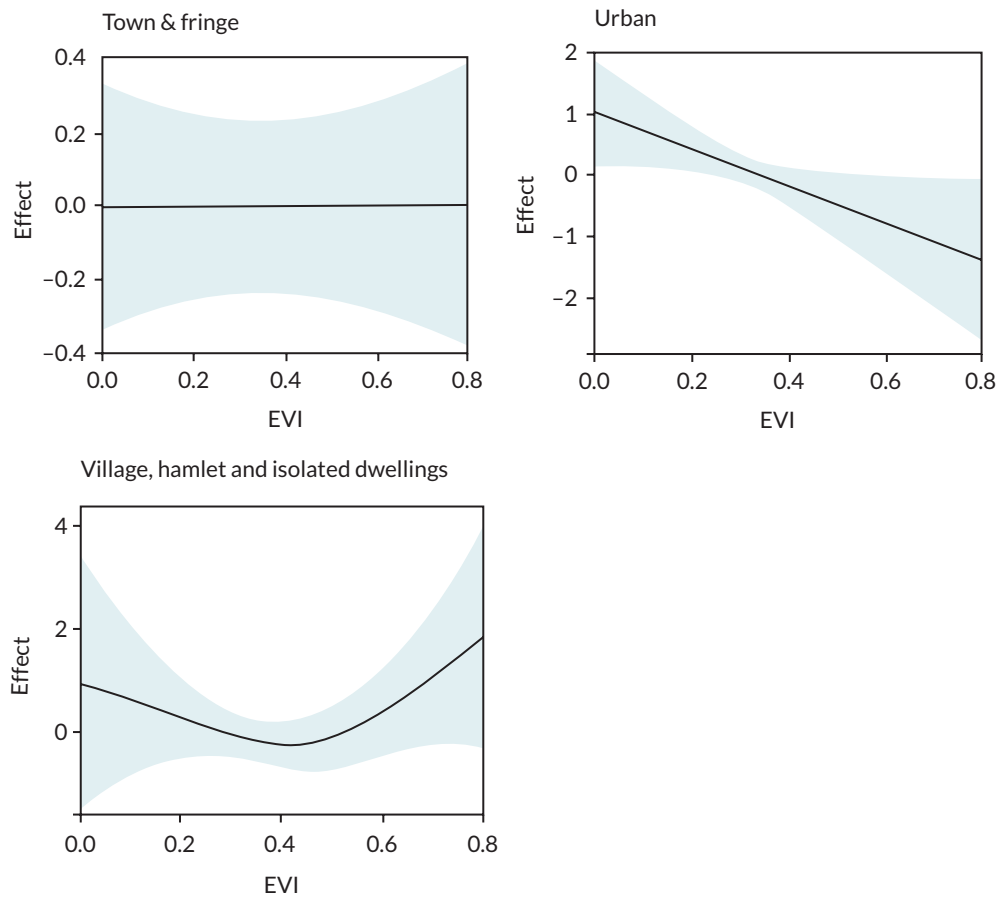
FIGURE 47 The smoothed function of EVI on WEMWBS by urban status. Models include WIMD, sex, age, economic status, material deprivation, car use, season and wave.
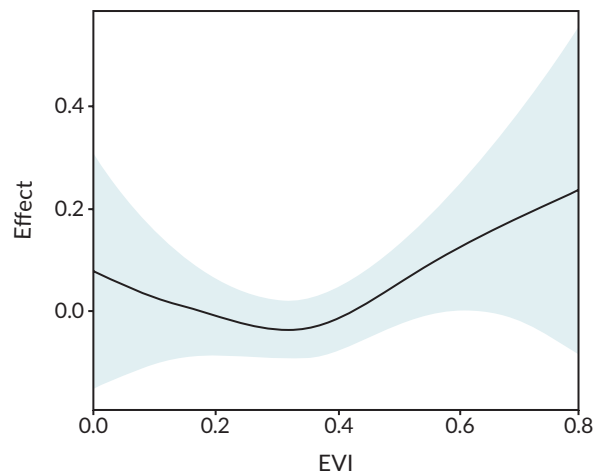
FIGURE 48 The smoothed function of EVI and the effect on life satisfaction. Model includes WIMD, sex, age, economic status, material deprivation, car use, season and wave.
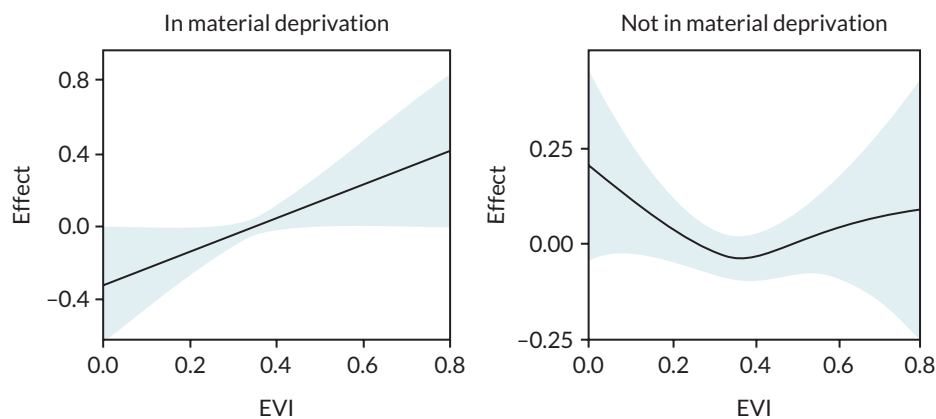
FIGURE 49 The smoothed function of EVI and the effect on life satisfaction by deprivation status. Model includes WIMD, sex, age, economic status, material deprivation, car use, season and wave.

Enhanced vegetation index predicting common mental health disorders

The CMD measure is binary, as such a quasibinomial distribution was applied in the GAM. EVI was not found to be related to CMD status for the full sample (edf < 0.01, p = 0.753). For those living in urban areas, EVI was significantly related but the relationship was 'wiggly' (edf = 6.36, p < 0.001). EVI was not significantly related to CMD for those living in town and fringe or village, hamlet and isolated dwellings (Town and fringe: edf = 0.45, p = 0.237; village, hamlet and isolated: edf = 0.01, p = 0.56). For those not in material deprivation, the relationship was approximating to linear (edf = 0.76, p = 0.077) and there was no relationship for those in material deprivation (edf < 0.01, p = 0.847).

Generalised linear modelling decision
Given the potential for a complex relationship between EVI and CMD, EVI was modelled as categories when predicting CMD. Given the high uncertainty beyond EVI = 0.6 (*Figure 50*) and the low number of people with a CMD where EVI > 0.6 (n = 26). The categories were < 0.2, 0.2–0.4 and > 0.4.

Proximity to nearest green or blue space predicting Warwick–Edinburgh Mental Well-being Scale

Proximity to nearest green or blue space was not found to be related to WEMWBS (edf = 0.02, p = 0.452). By urban status, proximity was not related to WEMWBS for those in either urban areas or town and fringe (urban: edf = 0.64, p = 0.09621; town and fringe: edf = 0.07, p = 0.358). For those living in village, hamlets and isolated dwellings the GAM indicated a complex relationship between proximity and WEMWBS (edf = 4.11, p = 0.009, *Figure 7*). By material deprivation, for those not in material deprivation proximity to nearest GBS was not related (edf = 0.02, p = 0.404) and for those in material deprivation the relationship was complex and marginally significant (edf = 3.94, p = 0.051).[268]

Generalised linear modelling decision
Categories were used in further GLM analyses which we categorised based on visually inspecting the relationships for both identified complex relationships (*Figure 51*) and adjusting cut-off points to the nearest 50 m. These were 0 to < 100, 100 to < 300, 300 to < 500 and > 500 m.
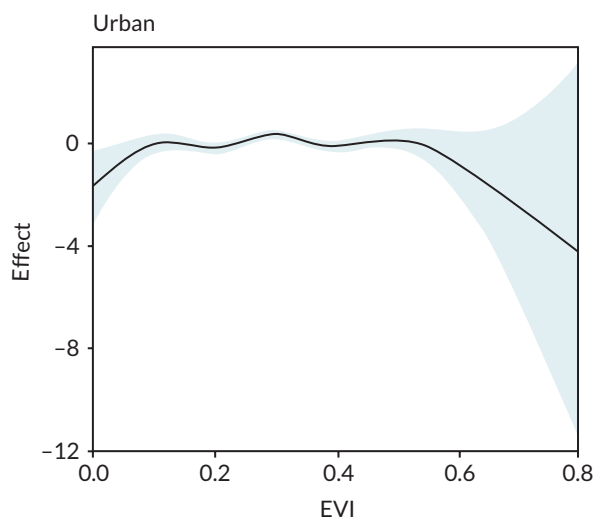
FIGURE 50 The smoothed function of EVI and the effect on life satisfaction by urban status (only Urban displayed). Model includes WIMD, gender, age, economic status, material deprivation, car use, season and wave.
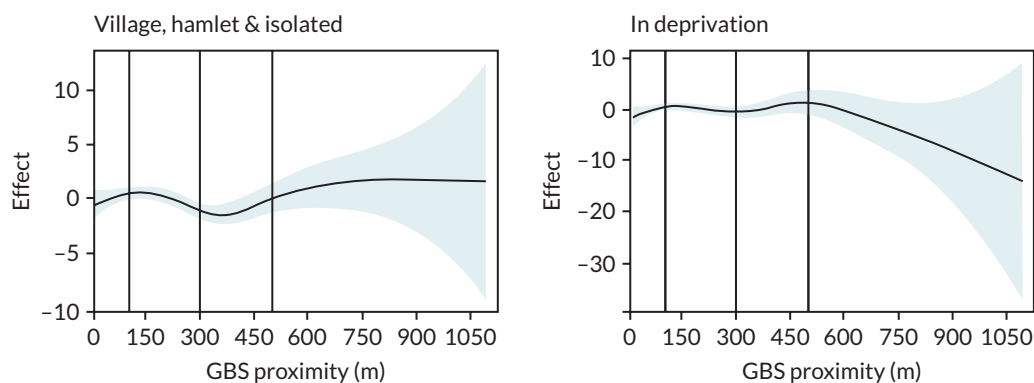

FIGURE 51 The smoothed function of proximity to nearest GBS and the effect on WEMWBS by: (a) urban status (only village, hamlet and isolated dwellings displayed); and (b) deprivation status (only in deprivation displayed). Model includes WIMD, sex, age, economic status, material deprivation, car use, season, wave and local authority. Vertical lines indicate the categories used in further analyses.

Proximity to nearest green or blue space predicting life satisfaction

Proximity to nearest green or blue space from the home was not found to be related to life satisfaction (edf = 0.02, p = 0.498). With separate relationships calculated by urban status, there was a significant, approximately quadratic, relationship between proximity and life satisfaction for those in living in town and fringe (edf = 2.11, p = 0.041; *Figure 52*). There was a marginally significant negative relationship for those in village, hamlet and isolated dwellings (edf = 0.66, p = 0.085) and no relationship for those in urban areas (edf = 0.01, p = 0.702). For those not in deprivation, proximity to GBS was not related to life satisfaction (edf = 0.01, p = 0.639) while for those in deprivation, the relationship was significant and complex (edf = 5.42, $p < 0.001$; *Figure 52*).

Generalised linear modelling decision

Given the complex nature of the relationship between GBS proximity and life satisfaction for those in deprivation and a preference for maintaining consistency between models, the same categories as used in modelling GBS proximity and WEMWBS were used in further GLM analyses for life satisfaction.
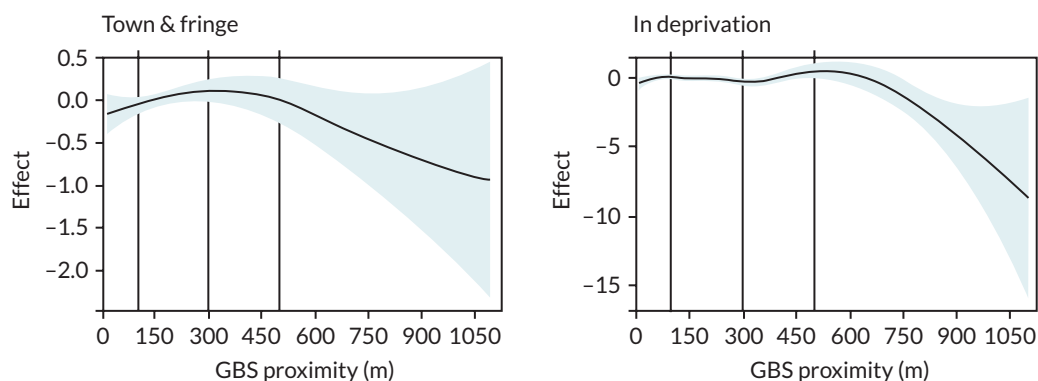
FIGURE 52 The smoothed function of proximity to nearest GBS and the effect on life satisfaction by: (a) urban status (only town and fringe displayed) and (b) deprivation status (only in deprivation displayed). Model includes WIMD, sex, age, economic status, material deprivation, car use, season, wave and local authority. Vertical lines indicate the categories used in further analyses.

Proximity to nearest green or blue space predicting common mental health disorder status

Proximity to nearest GBS was not related to CMD status (edf = 0.27, p = 0.299). When an interaction with urban status was included, proximity to nearest GBS was marginally significantly related to CMD status with an approximate quadratic relationship for those in urban areas (edf = 1.99, p = 0.076; *Figure 53*). For those in town and fringe and village, hamlet and isolated dwellings were not significantly related (town and fringe, edf = 0.00, p = 0.448; village, hamlet and isolated, edf = 0.00, p = 0.685). When an interaction with deprivation status was included, proximity to nearest GBS was marginally significantly related to CMD status for those in material deprivation with an approximate quadratic relationship (edf = 2.14, p = 0.066; *Figure 53*) while for those not in material deprivation proximity to nearest GBS was not related (edf = 0.00, p = 0.863).

Generalised linear modelling decision
Proximity to nearest GBS from the home was modelled as a linear and quadratic term in further GLM analyses.

Mediator – time outdoors

EVI is approximately linearly significantly positively related to time outdoors (edf = 1.04, p < 0.001) while proximity to nearest GBS was not related to time outdoors (edf = 0.07, p = 0.328; *Figure 54*). Time outdoors was significantly related to WEMWBS with a quadratic relationship (edf = 2.00, p < 0.001; *Figure 55*) and linearly significantly related to life satisfaction (edf = 1.09, p < 0.001; *Figure 55*). Time outdoors was significantly approximately negatively linear related to CMD status (edf = 0.92, p = 0.008; *Figure 56*). As time outdoors increases, the likelihood of a CMD decreases.

Modelling decisions
Time outdoors was modelled as a linear and quadratic term in further GLM analyses predicting WEMWBS and as a linear term in further GLM analyses predicting life satisfaction.

Mediator: time walking

EVI has a complex significant relationship with time spent walking (edf = 3.58, p < 0.001; *Figure 57*). Proximity to nearest GBS was not significantly related to time walking (edf = 1.00, p 0.287; *Figure 57*). There is a significant linear relationship between time spent walking and WEMWBS (edf = 1.00, p < 0.001; *Figure 58*). and a complex relationship with life satisfaction (edf = 3.45, p < 0.001; *Figure 58*).
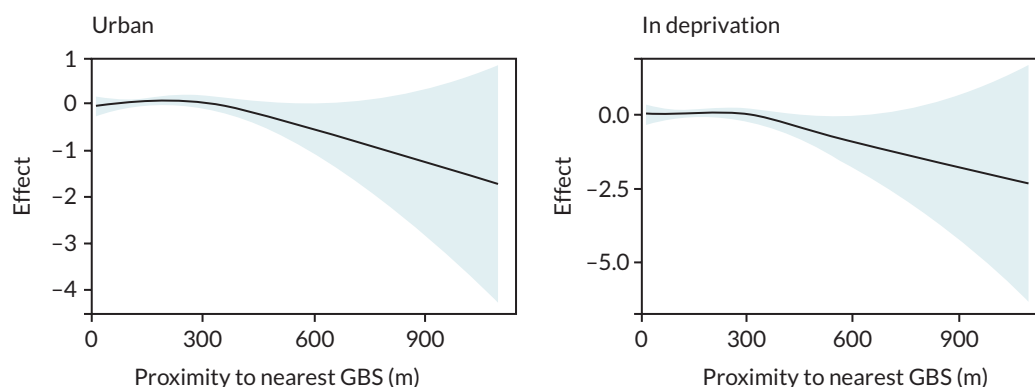
FIGURE 53 The smoothed function of proximity to nearest GBS and the effect on CMD status by: (a) urban status (only urban displayed); and (b) deprivation status (only in deprivation displayed). Model includes WIMD, sex, age, economic status, material deprivation, car use, season, wave and local authority.
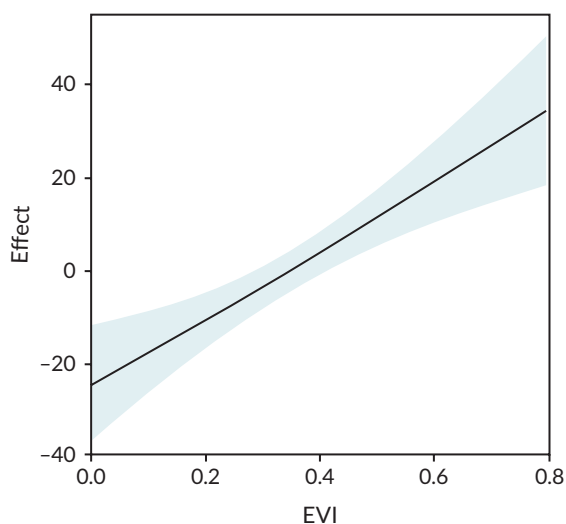

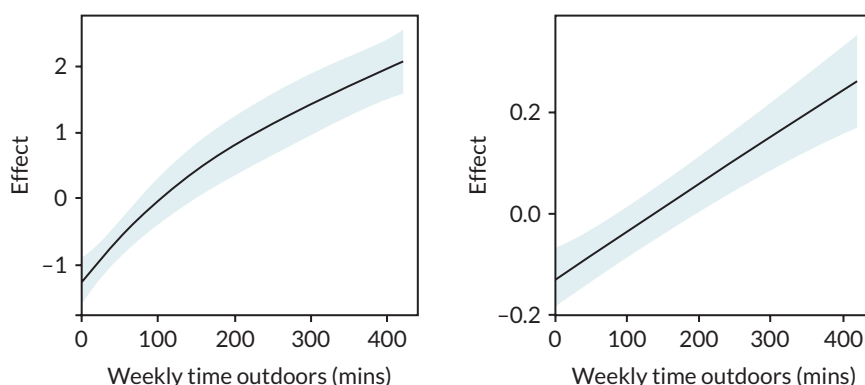**FIGURE 54** The effect of EVI associated with weekly time outdoors (minutes) (GAM).

FIGURE 55 The effect of weekly time outdoors as modelled with a GAM associated with (a) WEMWBS and (b) life satisfaction.
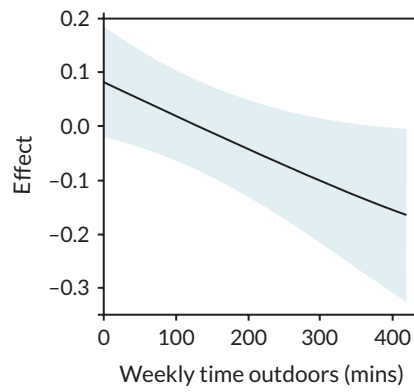
FIGURE 56 The effect of weekly time outdoors as modelled with a GAM associated with CMD status.
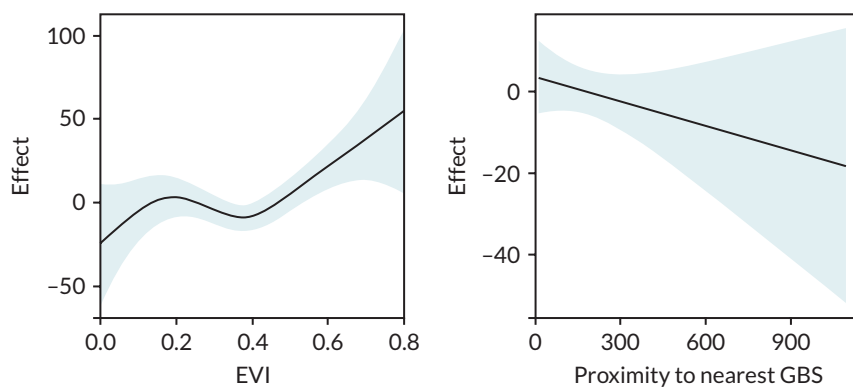
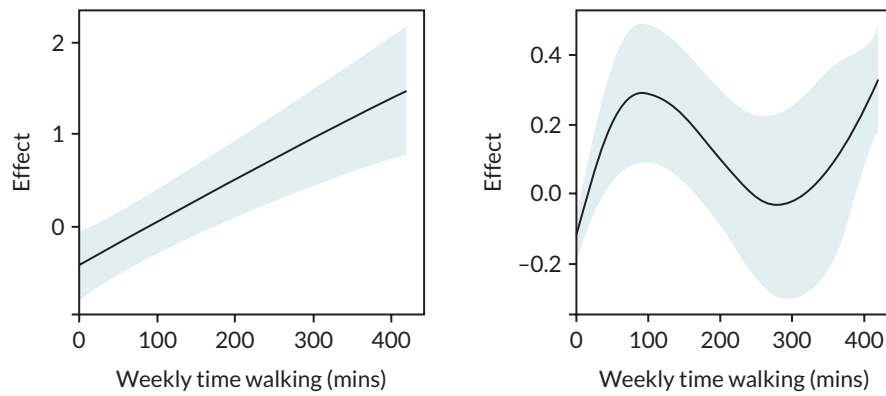**FIGURE 57** The effect on time spent walking by EVI and proximity to nearest green or blue space (GAMs).

FIGURE 58 The effect of weekly time walking (minutes) on WEMWBS and life satisfaction (GAM).

EME
HSDR
HTA
PGfAR
PHR