PHR Protocol – 16/07/07

Version: 0.1

Date: 24th November 2017

Green-blue space exposure changes and impact on individual-level wellbeing and mental health: a population-wide record-linked natural experiment

Chief investigator: Sarah Rodgers

Sponsor Swansea University

Funder Public Health Research programme

NIHR Portfolio number: PHR - 16/07/07

National Research Register: researchregistry2969

1. BACKGROUND

We will examine whether changes in subjective wellbeing, and the risk of common mental health disorders are associated with longitudinal time varying exposure to, and the use of, green-blue spaces (GBS). The evidence base on GBS impacts on mental health and wellbeing has been growing rapidly, with many studies suggesting positive relationships via mechanisms including stress recovery, cognitive improvement and social contact.(1–4)

A recent systematic review of the mental health benefits of long-term exposure to GBS in residential neighbourhoods found only "limited evidence" for a causal relationship between neighbourhood greenness and adult mental health, with key limitations being the limited number of studies and heterogeneity of exposure assessment; also, only 6 of the 28 studies included were longitudinal.(5)

Our approach will address the gap identified by this systematic review by providing further evidence regarding the association between health and GBS, considering both residential neighbourhood GBS exposure (longitudinally) and self-reported recreational visits to GBS (a hybrid cross-section/longitudinal design). Longitudinal measures will allow us to factor in individual-level loss and enhancement of residential GBS exposure. The national scale (Wales) of the proposed study permits *a priori* subgroup analyses to be planned (e.g. between those who have and have not moved) with sufficient power to detect clinically significant changes in our health outcomes. Our previous work has shown that about 10% of population move yearly.(6,7) Our statistical modelling will explicitly consider differential associations for key subgroups (e.g. deprived populations)(3) minimising risks that recommendations could have the potential to increase inequalities. Finally, we propose to model effect sizes by differential household level exposures, including a variety of GBS characteristics, for example size and type of GBS, providing evidence to inform decisions on the provision of GBS to town planners, local and central government.

2. RATIONALE

The research will estimate the impact of neighbourhood GBS exposure and visits to GBS on the risk of common mental health conditions, and subjective wellbeing; priorities for public health.(8) Common mental health conditions (including anxiety/depression) are experienced by around 1 in 4 of the UK population, and mental ill-health costs the UK economy over £100bn pa in health, social care and quality of life-loss costs.(8,9) With increasing impacts on wider societal costs, these issues are growing in importance. Subjective wellbeing is important in itself as a key marker of quality of life,(10) and is also related to mental and physical health outcomes, including life expectancy.(11,12) Access and exposure to natural environments – considered here as 'green-blue spaces' (GBS) – may provide opportunities to support and promote good public mental health and wellbeing.(13,14) A number of relevant theoretical frameworks to understand how GBS effects health have been proposed, with varying degrees of supporting evidence, including stress recovery, cognitive benefit and physical activity.(13,15) Recent systematic reviews indicated positive relationships between mental health and wellbeing and living near/visiting/exercising in GBS.(16–18) Longitudinal studies indicated that individuals' mental health and subjective wellbeing were improved during times spent living in areas with more urban greenspace, or closer to the coast.(4,19) However. the evidence base on this topic is limited, with few longitudinal studies, on changes in GBS exposure, or with assessments of visits to GBS.(5) Previous studies also lack censoring for births, deaths and migration. The proposed research explicitly addresses these gaps in the evidence. The project will capitalise on the opportunities presented by linking routine health data to environmental data to investigate secondary health outcomes that have also been associated with GBS exposure, including levels of physical activity.(20)

Improved evidence on the impact of GBS, and changes over time, on health is required to inform planning and area regeneration in the UK, which shape the environments people are exposed to. Natural England set out relevant evidence requirements(21) and conclude that "Whilst environmental

interventions targeting mental health appear to be effective, many of the existing studies are small scale and do not follow people for long enough for us to understand how effective the activities are". Our project partners NRW similarly require evidence to inform their policy and decision-making. The Welsh Outdoor Recreation Survey shows that recreational visits to natural outdoor spaces are declining (88% in 2011 to 84% in 2014 of people who had taken a visit in the last four weeks).(22) This study will consider neighbourhood GBS and visits to these spaces, providing evidence on how physical environmental change influences recreation. Our analyses will inform decisions on the planning of GBS to estimate the potential impacts on wellbeing and CMD. The evaluation of the long-term effects of access to GBS is requested in this commissioned call and we will complete this evaluation using 11 years of environmental exposure and corresponding health outcomes recorded for the majority of adults in Wales, with relevance to the wider UK population. This study will produce evidence addressing gaps stated by the Welsh Government Environment Bill White Paper; NRW recognises that inter-disciplinary studies spanning the environmental and socio-economic gradient are increasingly important to inform their decision making.

The current study adopts the approach of a natural experiment. MRC guidance on natural experiments indicates that the approach is appropriate when i) there is expectation of significant health impacts, but uncertainty over size and nature of these impacts; ii) randomised controlled trials would be impractical or unethical; and iii) the intervention or its underlying principles are potentially replicable, scalable or generalisable.(23) The study meets these criteria in that i) while GBS-mental health effect sizes may be small, the exposure is widespread and population health impacts potentially significant;(4) ii) RCTs that experimentally allocate long-term residential neighbourhood GBS exposure to participants are impractical; iii) enhancing/reducing access to GBS is feasible at scale. The study lies within a developing context of natural experimental research on environmental exposures, recognising the need for a relatively complex approach that is sensitive to the nature of an intervention exposure that is not fully randomised, blinded or necessarily easily parsed into 'treatment'/'control' status.(24–26)

3. OBJECTIVES AND OUTCOME MEASURES

3.1. Objectives

Our overall aim is to create novel data linkages between environment and health data to enable quantification of the impact of exposure to GBS through time for a national population, providing robust evidence to enable the future implementation of evidence-based policy. This directly addresses the NIHR commissioning brief PHR 16/07 "Change in health outcomes associated with access to and use of outdoor green and blue space and long-term effects?". We will exploit data, methods and systems available in Wales that are not available in the rest of the UK. However, we see no reason why the analyses for Wales should not, in principle, be generalisable to the UK. Some policy changes may be context specific, but we believe we can offer the best available analyses given the available data and systems. We will achieve our aim by completing several objectives (measurable and time-bound milestones are provided in section 15):

[1] Create longitudinal, residential GBS exposure data for all homes in Wales using Ordnance Survey, Local Authority and remotely sensed satellite data.

[2] Create longitudinal environment, health and demographic individual-level data linkages between survey and routine data within the Secure Anonymised Information Linkage databank (SAIL).

[3] Create a 11-year dynamic cohort of individual-level longitudinal residential GBS exposure to address the commissioning brief: "*Do people with different GBS exposures through time have different associated risks of having a CMD?*"" then, "Is the association between changes in GBS exposure and CMD modified by multiple socio-physical variables, migration and socioeconomic disadvantage?"

[4] Utilise the data-linked National Survey for Wales to address the commissioning brief focus on use of outdoor spaces by modelling interactions to answer research question: "what is the impact of GBS residential exposure, modified by GBS use and multiple socio-physical modifiers, on wellbeing, and common mental health disorders (CMD)?"

[5] Translate results into policy recommendations for government, and disseminate health outcome and economic impact assessment results to academics, stakeholders and policy makers. We will have involvement from public, planning and policy, including local park groups, Welsh Government, Natural Resources Wales, Local Authorities and Keep Tidy Wales,

We will add valuable, high quality evidence to a limited base, considering changes in residential GBS, and frequency of self-reported recreational visits. Longitudinal measures will produce robust findings to factor in both loss and enhanced residential GBS exposure. The national scale permits *a priori* subgroup analyses retaining power to detect clinically significant changes in our health outcomes. Our modelling will consider differential associations for key subgroups (e.g. deprived populations) minimising risks that recommendations could increase inequalities. Finally, we will create exposure measures to include a variety of GBS characteristics, allowing planners to consider the configuration (size, function, quality) of the most beneficial mix of spaces (see Planned interventions).

3.2. Proposed outcome measures

Outcome measure include a measure of mental health and a measure of wellbeing as joint primary outcomes over both work packages. Work package 1 (WP1: objectives 1, 2, 3) will use routinely collected data in SAIL to examine the risk of CMD using longitudinal changes in exposure to neighbourhood GBS. Work package 2 (WP2: objective 4) has in-depth survey responses from the data-linked National Survey for Wales (NSW), collecting a representative sample of Welsh residents' self-reported wellbeing and GBS use.

3.2.1. **Primary outcome (WP1):** Change in counts of common mental health disorder (CMD) treatments for adults with CMD recorded for the 78% of adults in Wales for whom we have GP data records in SAIL (1.7M adults).

3.2.2. **Primary outcome (WP2):** Subjective wellbeing (SWB) is our second primary outcome, measured by the Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS) in the NSW for two survey years (2016/17 and 2017/18) for a representative population (repeat cross-sectional survey, converted to hybrid longitudinal/cross-sectional design using data linkage variables). The WEMWBS is comprehensive (incorporating elements of both subjective and psychological wellbeing), short enough to be used in population-level surveys, responsive to change and has been validated among community samples of adults in the UK.(27) WEMWBS scores are on continuous scale from 14 to 70.

3.2.3. **Secondary Outcome:** Cost of GP events.;

3.2.4. **Secondary Outcome:** ONS-4 Subjective Wellbeing questions (https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing). The 4 ONS SWB questions are validated and used widely in national and international (OECD) wellbeing

surveys. Each of the four questions is on an 11-point (0-10) scale).

3.3. Preparation of datasets for measurement of outcomes

3.3.1. **Primary Outcome – Common Mental Health Disorder (CMD) (WP1)**

We used the same CMD algorithm in our earlier NIHR-PHR-funded study investigating changes in mental health as a result of a housing intervention. We used an algorithm from previous work(28,29), which was derived to estimate prevalence of CMD based on diagnoses, symptoms, and treatments validated against survey data. The authors compared routine data results to those from an MHI-5 survey with a cut point of less than 60 to allocate CMD case assignment. This work suggested several different algorithms providing different levels of sensitivity and specificity; the one we chose had the largest sensitivity without compromising on specificity.(28,29)

Common to all algorithms was the requirement of a diagnosis or symptom of a common mental health disorder prior to counting treatments. This is because several treatments have multiple purposes other than that of mental health disorders (e.g. antidepressants for pain treatment). On the advice of the authors for a previous (housing intervention) study we excluded the more severe condition of psychosis from these 'common' disorders; we will do the same for our proposed study to have maximum potential to detect change due to GBS exposure. The authors concluded their algorithm may be used to create outcome measures for trials and cohort studies. To have usability for the assessment of longitudinal changes we will count distinct relevant prescriptions per day, aggregated into quarterly time intervals for the CMD cohort. This will create individual level CMD counts for analysis. Further refinements to this method will be considered as part of this study; including adding prescription dosage, treatment type and quantities.

3.3.2. **Primary Outcome – Subjective Wellbeing (SWB)(WP2)**

The Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS) is a measure of mental wellbeing in the general population. We will model the WEMWBS as a continuous outcome. The WEMWBS has been well validated, with good internal consistency and reliability, and no ceiling or floor effects in a population sample.(27)

3.3.3. **Secondary outcome – GP events.**

We have the total number of events recorded for each person in a GP dataset. We will calculate the number of GP events, converting to distinct events. This eliminates double counting (e.g. counting the return of large quantities of test results only once).

3.3.4. **Secondary outcome – ONS4 Subjective Wellbeing questions.**

We will have individual responses to the 4 standard SWB questions as used by ONS and OECD. As well as being relatable to national representative survey data, these data have also been used in our previous green-blue space visit research using the comparable Monitor of Engagement with the Natural Environment for England (30).

4. RESEARCH DESIGN

We will evaluate the association between changes in GBS exposure, on the risk of CMD, and on subjective wellbeing. We will use cross-sectional and longitudinal data to examine the association between natural variability in time and space of GBS exposure and estimate whether this could be due to policies. We will overcome shortcomings of previous studies by creating a large time-varying

longitudinal GBS exposure dataset that will enable more precise estimates than had been achieved in previous studies, as well as estimating differential effects in particular population subgroups (e.g. by area level deprivation). We have developed a logic model, modified from Hartig and colleagues' (13) framework to inform the design of the proposed study, research objectives and analysis (Figure 1).

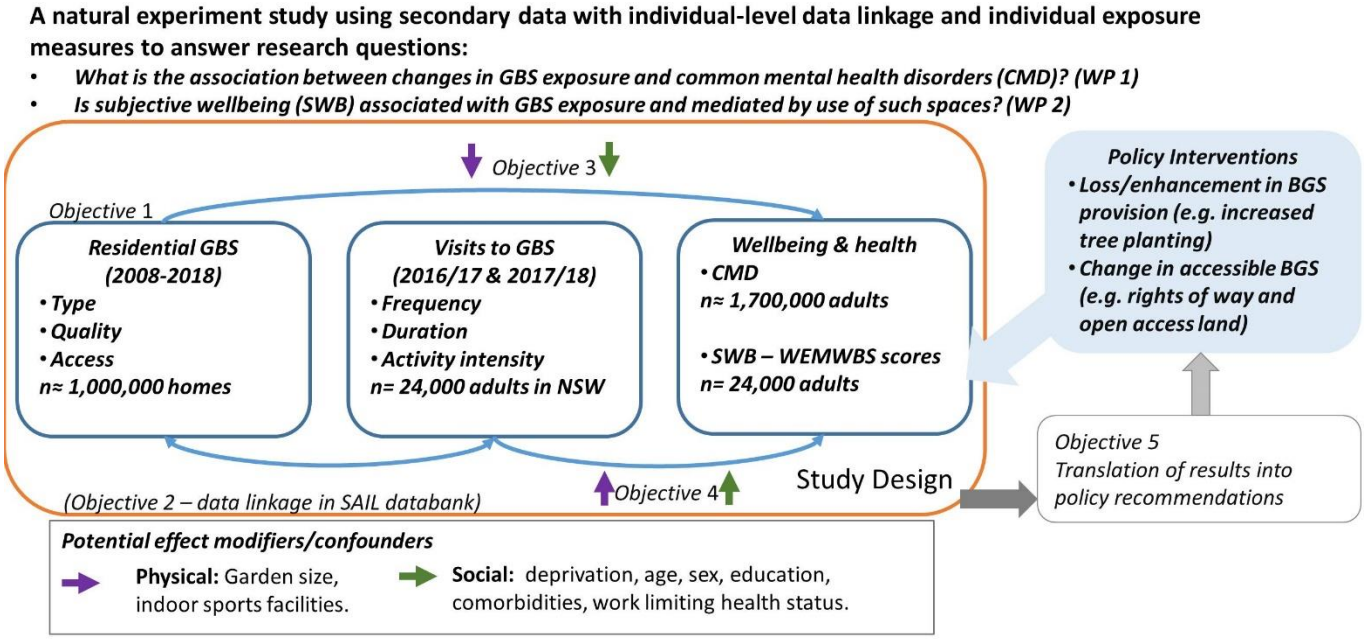

Figure 1: Logic model design. Diagram modified after Hartig T, et al. 2014. *Nature and Health*. Annu Rev. PH, 35:207-28

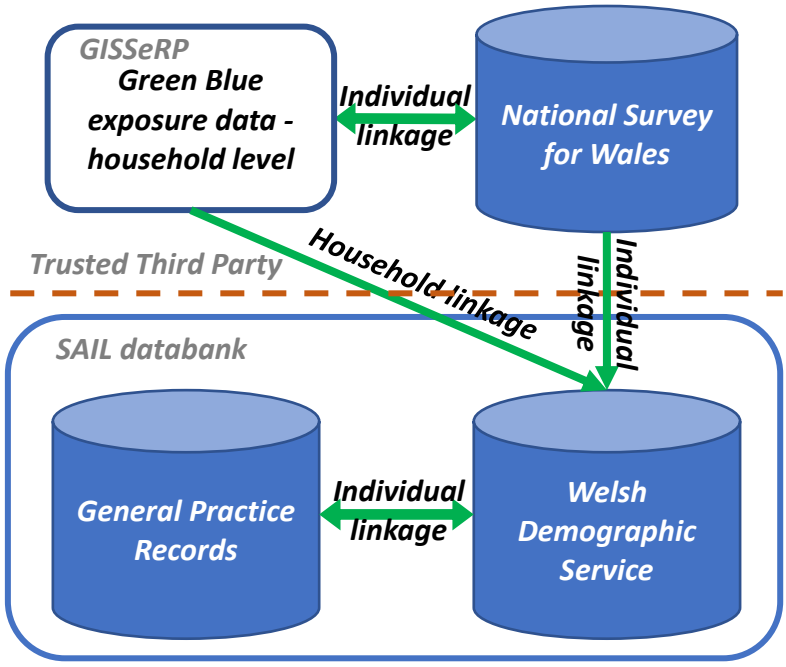

Figure 2: Data linkage diagram of SAIL databank. Demographics from Welsh Demographic Service (WDS), Common mental health conditions from GP records, GBS exposures created within GISSeRP, linked at household level, National Survey for Wales (NSW) linked at individual level by our Trusted Third Party (NHS Wales Informatics Service – (NWIS)).

4.1. Use of existing record linked datasets: The Secure Anonymised Information Linkage (SAIL) databank

The SAIL databank is held and managed at Swansea University and contains health, social and education data on three million residents of Wales, UK. It currently includes 150+ datasets and 50 billion records.(31–33) Information governance is overseen by an independent Information Governance Review Panel. Figure 2 illustrates data and linkages to be used in this project for both work packages.

4.1.1. Residential Anonymous Linking Fields

A strength of the SAIL system is the method for anonymising all households in Wales and linking household-level data from local authorities and others with individual health related data whilst protecting anonymity, using individual and residence-level linking fields. Address data are matched at NHS Wales Informatics (NWIS), where identifiable addresses are replaced with Residential Anonymous Linking Fields (RALF).(34) The residence-based metrics are then fully incorporated into the SAIL databank using residence to individual Anonymous Linking Fields (ALF). An environment Geographic Information System (see below) is used to create high resolution spatial metrics surrounding each residence. Over three million individuals whose data are held within SAIL can have made-to-measure environment data available to supplement their health, social and education data.

4.1.2. Welsh Demographic Service (WDS)

The Welsh Demographic Service dataset held by NHS Wales Informatics Service (NWIS), the NHS organisation in Wales mandated to hold personally identifiable data, contains addresses for all individuals who register with a General Practitioner. Dates for each address record update are held, thereby providing durations of residency for multiple homes and the ability to link to local environment exposures at each time point. This dataset holds demographic data including age and gender. These data may be used to create population sub-groups based on age, gender and location for each period. The WDS contains address information linked anonymously at the individual level (the ALF) which is the primary key variable for record-linkage. Using a split-file technique, NWIS supplies ALFs for the whole population of Wales to the SAIL databank.(31,32)

4.2. Geographic Information System (GIS) Secure e-Research Platform (GISSeRP)

Separate from the SAIL databank, is the Geographic Information System Secure e-Research Platform (GISSeRP) which contains map data from the Ordnance Survey at high spatial resolution. Ordnance Survey Master Map (OSMM) AddressBase Premium (ABP)(35) has point data for all residences and the OSMM Integrated Transportation Network dataset.(36) The ABP point data layer contains a point for current and historical residences, which is placed within the footprint of the residence. The buildings are surveyed with a spatial accuracy of ± 1-2 m. This provides geo-references at a high spatial resolution. The development of Green-blue space (GBS) exposure metrics are detailed in: 7. Planned Interventions. High resolution household identifiers (Unique Property Reference Numbers) are attached to each exposure metric, which are then used for anonymised linking into the SAIL databank. They are used for linkage and then removed, leaving linkages but no identifiers at the household level to preserve anonymity of the data.

4.3. Work Packages

WP1 (objectives 1, 2, 3) will use routinely collected data in SAIL to examine the risk of CMD using longitudinal changes in exposure to neighbourhood GBS in residents who have and have not moved home to distinguish between changes in GBS exposure as a result of moving home, and those that are a result of changes to land use, housing, access, policy (see logic model, Figure 1). We will use the Welsh Demographic Service (WDS) to calculate longitudinal changes in an individual's residential GBS exposure, both because of natural changes in GBS surrounding the home, and due to house moves. We have used these methods in our other NIHR PHR studies(37,38). Our spatial models of residential GBS exposure (rather than in small areas) allows us to calculate changes in their exposure for moves both between and within small areas (see planned interventions). A quarterly continuous measure of individuals' exposure to GBS will be accounted for using a panel design analysis. More details are given in 12. Statistical Analysis.

WP2 (objective 4) will use in-depth survey responses from the data-linked National Survey for Wales (NSW), collecting a representative sample of Welsh residents' self-reported wellbeing and GBS use.

Subjective wellbeing (SWB) is measured via the Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS), and the survey also provides detailed responses from about 12,000 people on their use of natural outdoor spaces (NSW: 2016/17, 2017/18). NSW questions on visits to natural spaces were part of the Welsh Outdoor Recreation Survey (2008, 2011 and 2014). We also have these earlier outdoor visit survey data to provide us with historical trends and descriptive contextual data. The 12,000 NSW responses will be linked at the individual level to GP and hospital admission data in SAIL, to examine associations between GBS exposure and activities with SWB and risk of CMD. We will use multilevel models to answer our research questions (see 12. Statistical Analysis). WP2 uses cross-sectional survey data, but adding longitudinal health and residential GBS exposure data is a hybrid cross-sectional/longitudinal design, permitting analysis of residential neighbourhood environment data as well as self-reported visit data to allow us to estimate direct and indirect effects.

4.4. Economic impact assessment

We will model the impact of GBS exposure on General Practitioner events and associated NHS costs. This is not a full economic evaluation but on the advice of a health economics expert (Hollingworth, Bristol) an impact assessment will add to the overall impact of our research. If we find a relationship between GBS, SWB and CMD, then we consider developing a method to extrapolate to estimate the impact of GBS on wellbeing for the national population. A reduction in wellbeing over time due to loss of GBS exposure could lead to increased healthcare utilisation and associated costs. Results will be generalisable to the other constituent countries of the UK and developed countries globally. This will form part of our reporting results objective [5].

5. RESEARCH SETTING

People aged 16 years and older living in Wales, UK. Dates for each update of the address record are held in SAIL, providing durations of residency for each successive address. This allows accurate exposures to neighbourhood GBS to be calculated for analyses. SAIL contains GP data for >70% of the population of Wales. People followed in the SAIL databank may move out of Wales or to a GP practice that does not provide data to SAIL, and are therefore lost to follow-up. We will exclude these missing periods from the panel analysis (see statistical analysis), considering each person's potential to accumulate health events down to a number of days within each quarter; using multilevel models approach in an unbalanced analysis does not depend on having equal number of participants for each period. Therefore, all adults with any duration of GP data will be included and retained for analysis; there will be no exclusions. We will have ~64,000 exclusions due to missing data linkage details (c.2% of total population). To maximise data utility, we will consider both imputation and the definition of 'missing' categories for each variable. We could then include the 'missing' categories in our models, treating the resulting regression coefficients as nuisance parameters.

6. PARTICIPANT ELIGIBILITY CRITERIA

People aged 16 years and older living in Wales, UK.

WP1) 22 local authorities, about 1.4 million homes, and 1.7 million adults (≥16 years) with GP data. To be allocated into groups according to change in GBS exposure, both people moving home and in situ environment changes; compared to a reference group without GBS change.

WP2) The representative sample of 24,000 adult participants of the National Survey for Wales will be linked into SAIL. Permissions for data linkage for the NSW have been negotiated previously (data provider agreements attached); data will be anonymised at the individual level via our trusted third party, the NHS Wales Informatics Service.(31)

7. Socioeconomic position and inequalities:

Evidence has been available for at least a decade suggesting that the public health interventions most likely to reduce health inequalities were those that operate at a higher societal level than one-to-one clinical interventions.(39) The more deprived the individuals, the more events conspire against them obtaining excellent preventative services and faithfully following the advice required to achieve full health benefits.(40) Universally legislated measures, such as water fluoridation, reach everyone. There is some evidence that accessible natural environments might provide some degree of resilience against socio-economic health inequalities.(3,41) Therefore, policies such as a national minimum standard for access to green and blue spaces may be considered a universal measure and could reduce health inequalities. Our analyses will include a large proportion of the entire population of Wales and use area-level and household level deprivation to analyse an area level natural experiment that will include 'participants' from across society, including the most deprived sectors of the population. Both work packages will consider socio-economic health and wellbeing inequalities:

- WP1) The Welsh Index of Multiple Deprivation (WIMD) is estimated and published at LSOA level.(42) Using the WIMD it is simple to categorise each LSOA to a deprivation quintile (based on 4 cut-points and equal counts) for aggregate analyses and to use the WIMD score as a LSOA-level covariate in statistical models. To estimate the effect of change in GBS on health and wellbeing-related outcome inequalities we will stratify the analysis by deprivation quintiles and model the interaction between WIMD and change in the GBS exposure.
- WP2) The NSW includes a measure of household material deprivation. People answering a deprivation question positively will trigger further questions to assess the extent of their deprivation. We will use this as a measure of individual socioeconomic position for use as a covariate in the analyses. Welsh Government NSW statisticians have provided the algorithm for classifying this variable and will provide the final derived variable as part of the 2016/17 and 2017/18 NSW datasets incoming to SAIL databank. We will fit a two-way interaction in the GBS exposure and material deprivation.

8. Planned interventions:

We are not allocating interventions. Instead we will consider retrospectively, the variation in green-blue space (GBS) potential exposure using accessibility, measured as a result of changes in GBS surrounding homes, and for individuals moving home. In this section we describe how we will achieve objective [1], the creation of a GBS exposure dataset. This is how we will capture retrospectively changes to the residential environment as a result of various policy interventions.

8.1. Who has delivered the interventions?

Local authorities have implemented different policies and areas have experienced different development activities, resulting in natural non-random variability in green-blue space in location and time. This is a natural experiment of the health impacts of changes in potential GBS exposure and as such we will not deliver an intervention but will analyse environmental changes that have been produced as the result of previous policy.

Analyses will focus on health outcomes varying through time, with access and exposure to GBS. This will be either as a result of moving to a new house (both blue and greenness changes), or through changes in exposure to greenness due to land use/cover or access changes (e.g. new pathways and entry points). Land use/cover and environmental accessibility are influenced by Local & Welsh Government policy and delivery strategies. A variety of policy and delivery strategy changes over the last decade, including those of Natural Resources Wales, a project partner, may have influenced the health of the Welsh population. These policies include increases in forestation,(43) (e.g. Forestry Commission), variation in accessibility(44,45) via access rights (sometimes related to grants), and the creation of new urban parks, such as the West Rhyl Greenspace, typically determined and implemented by Local Authorities.(46) The result of these changes will be captured within our GBS exposure dataset.

To obtain this information we will request and collate existing data from Local Authorities and our project partners to form a new GBS exposure dataset. It will contain information on the size, function (e.g. outdoor lido pool, football pitch, botanic garden, coastal path, woodland walk, playground), and quality of the open spaces, to enable us to assess most beneficial mix of spaces. This is in addition to the distance to each GBS, considering access point detail. Local Authorities have a Welsh Government mandated dataset providing these data. We will collate the data, so it is of most use to answering our research questions and influencing policy, as agreed with project partners and the public.

Our project partners have data on changes in rights of access and increased tree planting interventions. In addition to the creating a spatial model based on residential density of GBS, we will also augment this with data flags at the household level to create measures of loss and enhancement in policy intensive areas. Included in this flagged dataset will be Green/Blue Flag status, promotional quality indicators. We will work to combine the quality measures that are available from Local Councils, Natural Resources Wales and Keep Wales Tidy. We have established that the data received from one local council graded quality from an ecology perspective, and we will explore opportunities for development of consistent quality measures available with local council data and NRW. We will derive three key measures: 1) an ecological-perspective quality measure and 2) green flag amenity-focused quality indicator and 3) water-quality, amenity and interpretation-focused blue flag indicators. We will assess the potential to flag households within these “intensive” development areas, gathering information and feedback from/to local authorities along with their OpenSpace datasets, and additional data from our project partners.

Our natural experimental design (WP1) will evaluate change in residential GBS exposure through time on our outcomes. WP2 will use a hybrid longitudinal/cross-sectional design to estimate associations between recreational natural space visits to outcomes, in the context of residential GBS. We will produce findings to underpin future interventions, for example estimating effect sizes for interventions aiming to increase recreational GBS visit frequency.(20)

Strengths of our exposure data in this study are:
1) weighted network buffers of different distances to access GBS tailored to each person from their home, addressing ecological fallacy issues and an overreliance on small area measures using land use databases that have been relied upon in previous research;(4) and

2) detailed green space measures (based on satellite imagery/land cover databases) with data changing over time. Variations in different green spaces will be categorised to capture quality, size and function (OpenSpace assessments), and Green and Blue Flag data (Keep Wales Tidy). These will result in different spatial GBS exposure datasets.

A further major strength of this study is:
3) creation and data linkage of physical environment confounders/effect modifiers (e.g. access to sports centres) in addition to socio-demographic confounders currently available in SAIL.

8.2. Spatial Modelling of GBS data

We will measure the accessibility of GBS for each household using a network distance buffer (with distance decay weighting) to model access to residential GBS (Figure 3). We will adapt our previous CHALICE methodology that modelled household exposure to alcohol outlets to model green and blue spaces.(37) This will allow us to precisely define the exposure for each household.

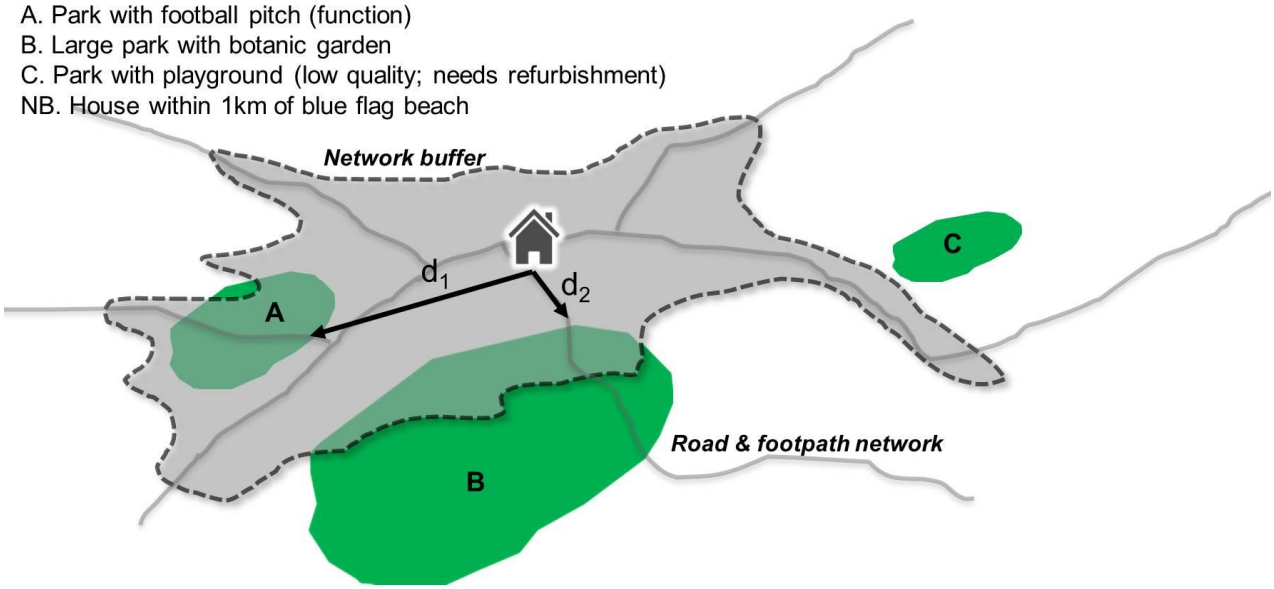

Figure 3. Residential GBS estimation for each home 1) Network buffer defines household exposure boundary (e.g. 10-minute walk). 2) GBS A and B included in exposure estimates (GBS C excluded). 3) Exposure measures include GBS quality, size, function and weighted network distances (d_i) to access points. 4) Longitudinal exposure estimates created through quarterly repeat analysis.

Change in GBS over our study period (2008-2018) will be created using temporally and spatially referenced satellite imagery.(47) Using imagery, we will extract temporal trends in of 'greenness' (48) for every household in Wales. These data will be used to create changes in individuals' GBS exposure scores, 2008 - 2018, to enhance the 2013 OpenSpace Assessment Data (collected under Welsh Government planning guidance)(49) and augment the all-Wales OS data we hold in our GIS database.(35,36,50) . We will also explore the integration of new green space and land use data as they become available and are published (e.g. OS Greenspace data (51)).

8.3. Control Treatment

All adults aged 16 years and over living in Wales will be included. Those who have moved out of Wales will be excluded from follow up in our study design that is capable of handling unbalanced data; that

is, their data from when they lived in Wales may still be included; only those people who experience a GBS change will have data that contributes to the model coefficients. The reference group will consist of those people whose GBS does not change. We are experienced in these designs following The Housing Regeneration and Health Study and CHALICE NIHR PHR studies.(37,52) Varying exposure to GBS will be calculated at a quarterly resolution prior to undergoing temporal treatments (see 12. Statistical analysis).

Our statistical model will use people who do not experience a different GBS exposure as the controls. This is a more robust method than using people as their own controls (difference in difference) and will use multiple variations in timing since GBS exposure change, with a maximum of 44 quarters possible over our 1 study period, without introducing dependence.(23) These will be people who do not move and for whom their residential GBS exposure does not change. This is the method we used in our recently completed NIHR PHR Housing study. The inclusion of many potential effect modifiers in our proposed study, together with more nuanced tailored GBS exposure data, has sufficient power to allow us to detect changes in either direction (loss/enhancement).(4)

8.4. Assessment and follow up

WP1: We will follow 1.7 million people for up to 11 years (2008-2018) in our dynamic natural-experiment longitudinal panel study using routinely recorded health outcomes (GP recorded diagnoses and prescription data). A subset of the population will have changed exposure due to moving home; another subset will have changed exposure due to changes in GBS in situ, (as assessed using the longitudinal GBS dataset collated from Ordnance Survey, Local Authority and remotely sensed data).

Access from people's homes to different GBS exposure will be evaluated within a natural experiment. GIS-generated GBS time-series variables will allow us to evaluate changes in local residential GBS for each individual. Exposure is defined by the Welsh Demographic Service (WDS) dataset (held within SAIL) providing precise beginning and end dates for residence of individuals in every household in Wales, enabling allocation of exposure of individuals to different types of GBS at a quarterly resolution, together with our outcomes extracted from routine data (e.g. CMD) in a longitudinal panel design. We will establish when the change in exposure began, changes in GBS exposure by type/size, and if there is a loss of GBS access, and if the individual moved out of Wales and can no longer be followed up. In a previous non-randomised regional intervention study (NIHR Carmarthenshire Housing) we had near-complete follow up for 183,553-person years; 45% of the 32,009 social residents were followed up for more than 10 years within a dynamic cohort. We anticipate similar follow up proportions for this proposed study but expanded to >70% of the national adult population in Wales (2,511,537 at baseline >= 16 years old (see Fig.1) x 70% = 1.7M) which is likely to result in the region of seventeen million person-years follow up, which will allow detection of short and long-term effects for population subgroups according to exposure and demographics.

WP2 uses NSW data comprising a different nationally representative sample in each survey and is a hybrid cross-sectional/longitudinal study. These people will be followed up for a shorter duration using both visit data, and time varying longitudinal residential GBS exposures, SWB and CMD health outcomes. The NSW data derive from a telephone survey carried out on behalf of the Welsh Government/Office for National Statistics.

9. STATISTICS AND DATA ANALYSIS

9.1. Proposed sample size:

WP1 primary outcome is prevalence of common mental health disorders (CMD). We will model the relationship between GBS and CMD as part of WP1. We completed pilot work using proxies for some of the variables due to lack of a fully developed household level GBS dataset (which required funding to complete). We analysed CMD diagnoses and prescriptions by extracting data from the SAIL databank in preparation for this proposal. We have used an approximation to the normal for the binomial distribution of prescriptions, comparing zero to more than zero prescriptions for a 1.5% difference in proportions and we have sufficient power (>90%) to determine whether there will be either loss of GBS, or enhancement. Our achievable sample size estimate is based on currently held SAIL data for a single year change in the small area environment domain, without an increase in income domain (Welsh Index of Multiple Deprivation). We require 23,265 individuals and we have 34,412 (see Figure 4, flow diagram) providing a comfortable margin and allowing the inclusion of several effect modifiers within interaction terms in the modelling process. While this currently ignores clustering within an LSOA, even if we include a correction factor of 2 for clustering, a common finding in work of this type, we would still have at least 80% power to detect a reduction of 1.5% difference in SWB proportions between the no change and low/enhancement of GBS. Additionally, our study will be 11 years long rather than a single year, providing over 90% power even with the inclusion of multiple effect modifiers. It is worth remembering that despite calculating this sample size using small area change data, the true numbers in this study will be based on data tailored to each individual's residential GBS exposure and therefore represents very conservative number for the actual study, for which the precision will be greater and even small differences in outcomes will be recognised.

Our proposed work will create a quarterly resolution dynamic cohort for 11 years of address history for all adults in Wales. We will also link to a high spatial resolution environment database. The completion of data linkage for one year ensured suitable sample sizes, and therefore the proposed study's viability (Figure 4). In preparation for this proposal we conducted some pilot work and grouped people who moved to a home in a different small area (LSOA) that had either the same or a different *environmental* domain score (loss, static, enhancement) according to the Welsh Index of Multiple Deprivation (WIMD). In our proposed study, we will link at the individual level; using data linkage to know if someone has moved home even within a small area and to potentially result in a different GBS exposure. Additionally, GBS surrounding people's homes has the potential to change using a dynamic environment dataset, and enhancement/loss will also be captured. For the outline proposal we filtered for adults only, calculated how many people moved home for one year (about 10%) and unlike previous studies that have used survey data, we will be able to follow up the majority of people, including having information on people who died. At outline stage, we included a proxy for a potential confounder by removing those people who had moved to a different *income* deprivation quintile, on the assumption that their circumstances may have changed radically, and for whom a change in GBS may be relatively unimportant. The inclusion of the WIMD income confounder makes this a very conservative estimate. A flow diagram of proposed study and numbers (1.7 million adults) we currently have before exposure group allocation is shown in Figure 5.

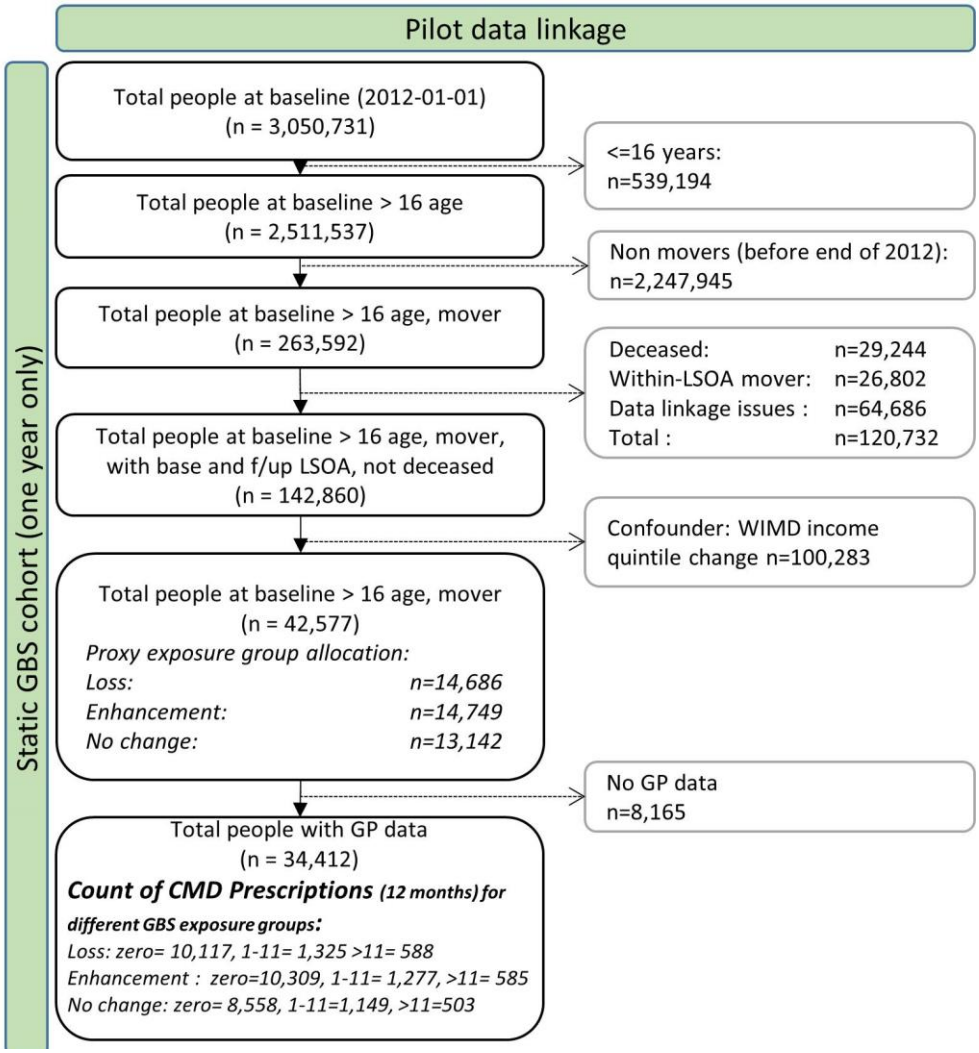

Figure 4: Flow diagram demonstrating data linkage of WDS and GP databases within SAIL databank for WP1. *Note: figures are for only one year and proposed work will use 11 years of dynamic cohort data (See Figure 5).*

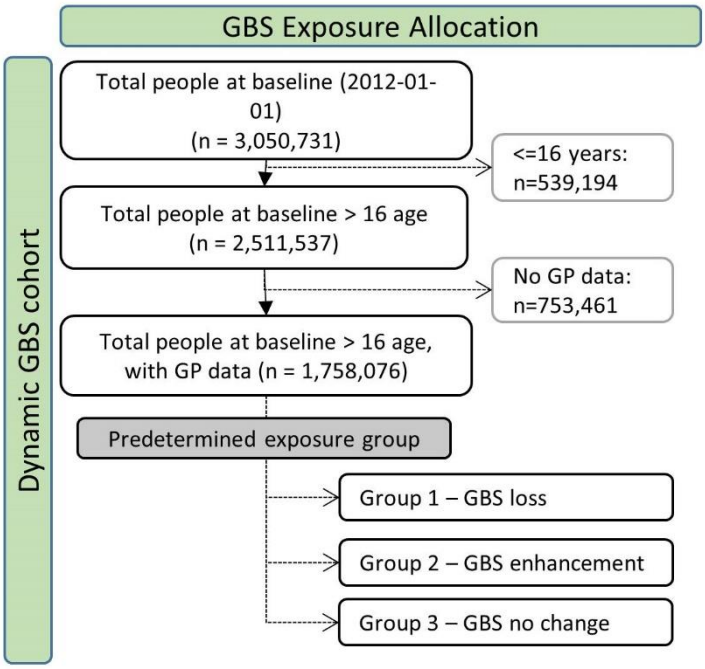

Figure 5: Flow diagram of data linkage of WDS and GP databases within SAIL databank for WP1 for dynamic cohort. Practices providing data to SAIL are increasing through time as more GPs are realising the benefit to their patients and practice. *Note: Individuals will be added to exposure groups once exposure datasets are complete.*

WP2: We have sufficient power to complete analyses in WP2. We will have 24,000 survey respondents data linked from NSW (2016/17 and 2017/18), of whom half (12,000) are asked about their visits to natural environments. This study will be powered at 90% to detect a mean change of 3 points on our WEMWBS primary outcome, requiring 468 participants.(53) Difference chosen as a minimum meaningful change, found in the evaluation of several different interventions. This will comfortably allow for clustering of GBS data and a reduction to 70% of this figure for including the GP data.

9.2. Statistical analysis:

The completion of longitudinal changes within panel designed datasets as part of objectives [1] and [2] will allow us to complete statistical analyses to answer the commissioning brief by completing objectives [3] and [4] and answering our research questions.

We will take into account inherent differences between individuals, or heterogeneity(54) by first adjusting for many individual level socio-physical variables,(55) (Table 1) and second by including an individual-specific random effect into the modelling framework(56), minimising potential bias of 'omitted effects' due to an individual's propensity to live in a home with a particular level of GBS access. For example, does interaction between individual's home-GBS and health vary significantly between people living in different deprivation areas?

Our analysis plan includes:

9.2.1. Descriptive Statistics

Full descriptive statistics of GBS exposure and changes in GBS exposure by household, LSOA, deprivation quintiles using WIMD and population characteristics (e.g. age, gender). National Survey for Wales data will have descriptive statistics of each of the survey questions we will use (number of visits, labour market status, etc.)

9.2.2. Temporal treatment of spatially modelled GBS exposure data

A spatial model of changes in GBS exposure at individual-level data will be created to explore the relationship between counts/rates of our outcome measures and changes in GBS exposure. The residential exposure models will be temporally treated to allow us to account for measures of change in exposure. We will use methods based on previous NIHR funded work by members of the research team.[3] We will model change in GBS exposure preceding outcome events: (a) the previous quarterly exposure; and (b) the change in exposure (the difference between quarterly values 1 year apart divided by the square root of the mean of the five quarterly values). At this stage, prior to data linkage, we will produce maps of the GBS spatial models (without health outcomes) as a valuable output for project partners to see where changes in GBS are occurring nationally. This will form part of our dissemination (objective 5) and may incorporate an online interactive map demonstrating changes over different periods.

9.2.3. Geographically Weighted Regression (GWR)

We will use GWR(57) models to explore the relationships between changes in GBS exposure and our primary and secondary outcomes as a function of changes in GBS score outlined above (2). GWR is slightly different from other spatial regression techniques such as the spatially lagged model and spatial error model in the sense that it enables the spatial localisation of standard regression models thereby, allowing us to incorporate the spatial heterogeneity of the degree of accessibility/exposure to GBS in explaining the spatial distribution of health outcomes. Different types of spatial models make different assumptions, model spatial autocorrelation differently and analyse the data in different ways; their relative merits will be explored as part of the project. We will use GWR as complementary analyses to the multi-level modelling set out in table 1, providing context to these results and exploring local non-stationarity at neighbourhood level in relation to GBS and our primary and secondary outcomes.

9.2.4. Mixture of outcome types

Across both work packages, we have outcomes of two distinct types: continuous (SWB) and counts (CMD and GP event days). Each type of outcome will require a different member of the family of generalized linear models. The continuous outcome will be analysed using a linear model. Poisson models will be used to analyse the longitudinal count data. In each work package, each of the two outcomes will be modelled separately. The count data will have an exposure variable, indicating the number of times the event could have been recorded in the data; that is, the period of data we hold for each person (down to number of days). This variable will be incorporated into a Poisson model.

9.2.5. Statistical modelling approach

Our data are hierarchical in nature: WP1: three level data: LSOA / Individual / Time (Quarter); WP2: two level data: LSOA / Individual. Accordingly, we will adopt a multi-level modelling approach. We will generalise the standard linear and Poisson models to handle two or three levels of variation, as appropriate. The resulting two- and three-level models will allow us to estimate the extent to which variation at each level may be explained by confounding variables (see Table 1). We will examine the

extent to which the effects of the exposure variables are mediated by this range of confounders. We will adjust for these confounders by including them in our models first. Having controlled for confounders, we will proceed to add the exposure variables to our models, thereby allowing us to test for the significance of the exposure variables in the presence of the confounders. Inclusion of these variables in the above models will permit us to estimate their direct effects on our outcomes.

9.2.6. Interactions

Having included the main effects of both confounders and exposure variables, we will proceed to test for the significance of selected pairwise interactions of interest, as indicated in the interactions column of Table 1.

Our statistical analysis plan is summarised in Table 1.

Table 1: Statistical Analysis Summary (P: Primary; S: Secondary)

WP	Outcomes (Type)	Exposures	Confounders	Interactions of interest
1	P1: CMD (count) S1: GP event days (count) (and cost)	P: Exposure to residential GBS. S: Timing of move(s); Time since last move	LSOA level: Quintile of deprivation (WIMD). Category of urban/rural settlement type (ONS classification). Car ownership variables from the census Individual level (SAIL): Gender, Age, Comorbidities (Charlson Index), Distance from nearest council & private sports centre, public transport accessibility to GBS. Quarterly level: Seasonality.	Change in GBS exposure by Deprivation Change of GBS exposure by size, quality and function
2	P2: SWB (continuous) P1: CMD (count) S2: SWB ONS4 (continuous)	P: Exact number of visits made outdoors for recreation in last 4 weeks S1: Level of engagement in 150+ minutes of moderate or vigorous intensity activity per week S2: Proportion of all physical activity which is undertaken in green-blue space P: Exposure to residential GBS.	LSOA level: as per WP1. Individual level (SAIL as per WP1) and NSW: Gender, Age, Highest educational qualification, Marital status (incl. living with partner). Residence type, Number of children living at home, Material deprivation, Labour market status, Work limiting health status. Distance from nearest council & private sports centre.	Exact number of visits by Deprivation Exact number of visits by access to sports centres Exact number of visits by Material Deprivation Exact number of visits by residential GBS

Controls are people who do not move, and for whom their GBS exposure does not change.

Research Question - objective 3: "Do people with different GBS exposures through time have different associated risks of having a CMD?" then, "Is the association between changes in GBS exposure and health modified by multiple socio-physical variables, migration and socioeconomic disadvantage?"

WP1: We will have multiple observations for everyone in a panel design multilevel model with time as the lowest clustering level. This will allow us to consider the inherent autocorrelation in repeated observations for the same individuals over time: 44 quarterly time periods will be nested within individuals, nested within small areas. This statistical model will allow us to answer our novel research question. We will primarily model exposure change but will retain a 'move flag' so we may refine our estimated effects by looking explicitly at subgroups such as: those who moved versus those people who remained in the same home but for whom their residential GBS has changed (loss/enhanced GBS), (see control within Planned Interventions).

Using the vast quantities of data within SAIL, we will assess the potential to use multilevel models to explore further interesting relationships: random intercept by random slope covariance could be used to unpack any non-linearity in the data. For example, for a change in GBS residential exposure in urban areas with little accessible green space may have a much stronger effect than the same change in green space in areas with already high levels of green space (following from the modification "Exact number of visits by residential GBS" (Table 1)).

Research Question - objective 4: "what is the impact of GBS residential exposure, modified by GBS use and multiple socio-physical modifiers, on wellbeing, and common mental health disorders (CMD)?"

WP2: will have individuals linked to their outdoor visit and physical activity (PA) data, clustered within small areas, allowing us to answer the question, as instigated by logic model (Figure 1): ***is there a main effect of association between visits to outdoor spaces and SWB?*** (See Table 1, WP2).

We will then progress to answer a further question combining relevant cross sections of visits with residential GBS exposure data from WP1 to answer a further question: ***Is there a main effect of association between residential GBS and SWB & CMD, and is this mediated by visits to outdoor spaces?*** As the logic model suggests, some pathways between residential GBS and health outcomes rely on actually visiting the spaces, whereas others work through more indirect mechanisms and will simply depend on residential proximity; that is they do not need to visit to benefit. It may be that benefits occur from just being able to see the space or even to know one is close. We will adjust for several potential confounders and effect modifiers, including many socio-physical factors (See Table 1). Mediation analysis will allow us to investigate to what extent residential GBS associations are explained through visit frequency (see next).

9.2.7. Mediation Analyses – Physical Activity (WP2)

We will include as a mediator, estimates of Physical Activity(PA) from the NSW that asks about moderate intensity physical activity equivalent to >3 Metabolic Equivalent of Tasks (METs), to quantify outdoor GBS PA.(58) Metabolic Equivalent of Task (MET) is a physiological measure expressing the energy cost (or calories) of physical activities. Three METs results in increased heart and breathing rate, etc. in a manner similar to other general population studies.(59) The measure of physical activity that will be used for this study will be meeting the current WHO Physical Activity standard of 150+ minutes of moderate or vigorous intensity activity per week, through participation in outdoor recreation in green-blue spaces. This measure will be determined through the analysis of a combination of questions from the new National Survey for Wales (NSW), specifically those relating to: (a) the exact number of visits made to the outdoors for recreation in the last 4 weeks, (b) the intensity of activity undertaken, (c) the time spent doing the activity. We will also derive a variable for *total* PA by crossing GBS PA with general exercise questions for vigorous and moderate activity: "*Think*

about the moderate physical activity you did in the last 7 days. Moderate activity takes some effort and can make you breathe somewhat harder than normal. Examples include digging in the garden, spring cleaning or other heavy housework, gentle swimming or cycling." This has been added due to our experience in analysing MENE data (Monitor of Engagement with the Natural Environment, Natural England) where 3.2M people who achieved 150mins pw, wholly or partly through green/blue exercise, only 43% achieved it *wholly* in GBS.(1,60) A further 57% achieved their recommended PA partially in GBS. We feel to leave out these 57% (1.8M people) would lead to an underestimate of the value of blue/green exercise for health. Weekly PA in the NSW is based on the International Physical Activity Questionnaire (IPAQ), an internationally standardised PA questionnaire, allowing us to precisely define physical activity levels. The use of >3 METS will enable us to cross-check our values with the WHO HEAT tool for walking, a tool widely used for evidencing policy changes. To augment other (indoor) PA, we will use the total *opportunities* for PA within their locality, for instance sports facilities within a kilometre of home. We will also use sport centre usage figures from the Active Adult survey (annual data 2012-2016), providing a representative sample density of national usage of sport centres and a trend over time.

9.2.8. Structural Equation Modelling (SEM)

To further explore mediation, we will consider developing a structural equation model based on: a) The "Hartig" theoretical model (Figure 1); and b) the results from the previous regression models exploring each of the stages in isolation. For example, does GBS where you live (residential) or use of GBS (visits) affect PA differently to result in different impacts on health outcomes? A SEM can statistically parse out the amount explained by a predictor variable, on any given outcome variable (the 'total effect'), in terms of the 'direct effect' and 'indirect effect' (causality should not be assumed). The direct effect refers to that amount of the association between the predictor and outcome variable which remains even when the 'mediator' is controlled for. The indirect effect refers to the amount of the association between the predictor and outcome variable that is affected by the mediator being introduced into the model, and thus presumably affected by it. A significant indirect effect is usually interpreted as a sign of potential mediation. In Figure 1, for instance, physical activity is conceived as one of several potential mediators between local area greenspace and physical health, i.e. *one of the reasons why people are healthier in greener areas is because they are more likely to engage in physical activity than those in less green areas.* A SEM approach will test this hypothesis by comparing a pre-specified model based on theory with the actual statistical data collected to see how well the theoretical model "fits" the data. SEM also provides suggestions for how a model could be improved, so called 'modification indices', which can shed insight into how the data are actually related, over and above the initial theoretical model.

10. Public Patient Involvement

We will invite members of the public to a workshop with investigators where we will present preliminary results and ask them to help direct our research through a series of focussed questions. The workshop group will comprise members who are experienced at considering the value of environment from the Health and Environment Public Engagement (HEPE) Group (hosted by University of Exeter), who will join Wales-based members of the public from urban park groups (e.g. friends of Brynmill Park in Swansea), and those experienced in considering data linkage proposals. HEPE are the PPI group and leading public engagement activities of the NIHR Health Protection Research Unit in Environmental Change and Health (which includes the Exeter-based investigators and Public Health England). We consider that this group of mixed backgrounds will make for an extremely valuable public engagement process, and will be essential in directing the second half of the project, along with results from previous research. We will also invite project partners, including NRW and Keep Wales Tidy. We will also consult this group on an ongoing basis to ask for their advice in the presentation of our results. HEPE group members have already been consulted during proposal

development, and have expressed support for the activities proposed, and provided input on the plain English summary.

We have project partner agreements from key stakeholders who will form a study steering committee including Local Authority planning and data experts (e.g. Merthyr Tydfil and Swansea), Natural Resources Wales, Keep Wales Tidy, Welsh Government and Sports Wales. Please see their attached letters of support. We will recruit members of the public to the study steering committee (SSC) from the DECIPHer Alpha group (aged 16 and over), the Welsh Government's Involving People Network, and with the help of the SAIL consumer panel. The benefit of having a SSC is that they will be able to advise the research team on the design, result interpretation and dissemination of the research findings to wider academic, policy and public audiences. Policy members of the SSC will advise on national and local level policies for green-blue space development and management; public members of the study steering committee will advise on how the public interact with and use green-blue space. The study steering committee will be reimbursed for their time using the NIHR INVOLVE guidelines. We have developed this proposal with project partners NRW. Linking with these networks will allow us to engage with a wider network of experts and interested parties to advise us and receive feedback on our proposed methods throughout the project lifetime. Moreover, the networks will help us to disseminate our findings to a wide audience of interested academics, policy stakeholders and the public.

Our project has been adopted by the Farr Institute and UKCRC Centre for Public Health Excellence (DECIPHer: Centre for the Development and Evaluation of Complex Interventions for Public Health ImpRovement). The project links strongly with activities of the NIHR Health Protection Research Unit in Environmental Change and Health (involving Exeter co-applicants, with Public Health England, Met Office, LSHTM, UCL), the Horizons 2020 'BlueHealth' project (led by the University of Exeter), and the FP7 Phenotype study.

11. ETHICAL AND REGULATORY CONSIDERATIONS

11.1. Ethical arrangements

This study is based on routinely collected administrative, environment, and survey data. All data will be anonymised into a databank (SAIL) and therefore there will be no mechanism for informing potential study participants of possible benefits and known risks. This is a retrospective evaluation of a natural experiment and therefore all changes to environments have already occurred as part of evolving changes due to policy and planning, or due to migration choices made by families or individuals. We have obtained informed consent to anonymise and link the National Survey for Wales data into the SAIL databank. All routinely collected anonymised data held in SAIL, including NSW, the hospital admissions in the Patient Episode Dataset for Wales and common mental health disorders in the GP dataset, are exempt from consent due to the anonymised nature of the databank (under section 251, NREC).

11.2. Research Governance

We are proposing the use of anonymised data and therefore we are exempt from NREC. We are in the process of applying to our independent Information Governance Review Panel (IGRP) for permission to conduct this study. We explain this in our ethics section. The IGRP contain independent members from NREC and BMA, as well as lay members, and have previously given permission for similar projects (e.g. NIHR PHR CHALICE and NIHR PHR Carmarthenshire Housing). The review process checks that the study we are proposing is useful, not service evaluation, and will not break anonymisation standards. The review normally takes about 6 weeks and we do not foresee any problems in the IGRP granting

permission to complete this study. We will have a Study Steering Committee (SSC) and have described partner collaboration membership in that relevant section, and justified resources. We will also invite key independent academics to join this committee. Swansea University will be the nominated sponsor.

12. DISSEMINATION POLICY

Throughout the course of this research we will report our progress to the study steering group to inform them of progress and receive feedback. The steering group will comprise stakeholders as detailed in PPI and we have included appropriate costs for their involvement. Findings from the study will be published in a comprehensive report and we will notify all relevant stakeholders and promote publication through networks established during this research. At the end of the study we will hold a workshop to report our findings to stakeholders and the public. We will disseminate our findings to patient, policy and academic networks. We will produce a plain English summary of the report to be distributed around Farr, ADRN and NIHR to disseminate the findings to wider networks and we will present results to the public with easily accessible media (infographics). All findings will be shared through our websites and social media to maximise international exposure. We will present findings via seminars to key health professionals, including Public Health England and Public Health Wales, health service commissioners, LGA and government planning officials to make recommendations for future policy decisions in this area and to those who have an interest in improving GBS, CMD and SWB. Further to this we will disseminate the report findings via links on PHW website to make sure the findings of the research reach relevant political stakeholders. Academically, we plan to publish papers in internationally peer reviewed journals to disseminate the research to interested individuals in the wider academic community and add to the evidence base. We will publish many papers such as: study protocol, methodological papers on GIS and statistical analysis, results papers on wellbeing and mental health. We will share our results at national and internationally recognised conferences and promote our findings in academic circles. All activities have been costed appropriately.

13. 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 Mar;10(3):913–35.
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 [Internet]. 2014 Apr [cited 2015 Sep 17];4(4):e004951. Available from: http://bmjopen.bmj.com/content/4/4/e004951
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 [Internet]. 2015 [cited 2016 Jul 28];14:17. Available from: http://dx.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 Apr;956797612464659.
5. 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 [Internet]. Multidisciplinary Digital Publishing Institute; 2015 Apr 22 [cited 2016 Jul 11];12(4):4354–79. Available from: http://www.mdpi.com/1660-4601/12/4/4354/
6. Hutchings HA, Evans A, Barnes P, Healy MA, James-Ellison M, Lyons RA, et al. Does frequent residential mobility in early years affect the uptake and timeliness of routine immunisations? {An} anonymised cohort study. Vaccine. 2016 Feb;
7. Hutchings HA, Evans A, Barnes P, Demmler J, Heaven M, Hyatt MA, et al. Do Children Who Move Home and School Frequently Have Poorer Educational Outcomes in Their Early Years at School? An Anonymised Cohort Study. Kemp AH, editor. PLoS One [Internet]. 2013 Aug 5 [cited 2016 Dec 9];8(8):e70601. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23940601
8. Davies S. Annual Report of the Chief Medical Officer 2013. Public Mental Health Priorities: Investing in the Evidence. London: Department of Health; 2013.
9. Economic and social costs of mental health problems [Internet]. Centre For Mental Health. [cited 2016 Jul 28]. Available from: https://www.centreformentalhealth.org.uk/economic-and-social-costs
10. The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. Soc Sci Med. 1995 Nov;41(10):1403–9.
11. Diener E, Chan MY. Happy People Live Longer: Subjective Well-Being Contributes to Health and Longevity. Appl Psychol Heal Well-Being. 2011 Mar;3(1):1–43.
12. Steptoe A, Deaton A, Stone AA. Subjective wellbeing, health, and ageing. Lancet. 2015 Feb;385(9968):640–8.
13. Hartig T, Mitchell R, de Vries S, Frumkin H. Nature and Health. Annu Rev Public Health [Internet]. 2014 Mar 18 [cited 2016 Jun 8];35(1):207–28. Available from: http://www.annualreviews.org/doi/abs/10.1146/annurev-publhealth-032013-182443
14. Marmot M. Fair society, healthy lives [Internet]. 2010. Available from: http://www.ucl.ac.uk/gheg/
15. Lachowycz K, Jones AP. Towards a better understanding of the relationship between greenspace and health: {Development} of a theoretical framework. Landsc Urban

Plan [Internet]. 2013 Oct [cited 2016 Jul 28];118:62–9. Available from: http://www.sciencedirect.com/science/article/pii/S0169204612002939
16. Bowler DE, Buyung-Ali LM, Knight TM, Pullin AS. A systematic review of evidence for the added benefits to health of exposure to natural environments. BMC Public Health [Internet]. 2010 [cited 2016 Aug 8];10:456. Available from: http://dx.doi.org/10.1186/1471-2458-10-456
17. Dzhambov AM, Dimitrova DD. Urban green spaces' effectiveness as a psychological buffer for the negative health impact of noise pollution: a systematic review. Noise Heal. 2014 Jun;16(70):157–65.
18. 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 Mar;45(5):1761–72.
19. White MP, Alcock I, Wheeler BW, Depledge MH. Coastal proximity, health and well-being: {Results} from a longitudinal panel survey. Heal Place. 2013 Sep;23:97–103.
20. Hunter RF, Christian H, Veitch J, Astell-Burt T, Hipp JA, Schipperijn J. The impact of interventions to promote physical activity in urban green space: a systematic review and recommendations for future research. Soc Sci Med. 2015 Jan;124:246–56.
21. England N. Links between natural environments and mental health: evidence briefing - {EIN}018 [Internet]. Natural England - Access to Evidence. [cited 2016 Jul 28]. Available from: http://publications.naturalengland.org.uk/publication/5748047200387072
22. Williams S. Wales Outdoor Recreation Survey 2014: Final Report [Internet]. 2015. Available from: http://www.ccgc.gov.uk/publications--research.aspx
23. Craig P, Gunnell D, Haw S, Lawson K, Macintyre S, Ogilvie D, et al. Using natural experiments to evaluate population health interventions: new MRC guidance. J Epidemiol Community Health. 2012 Dec;66(12):1182–6.
24. Humphreys DK, Panter J, Sahlqvist S, Goodman A, Ogilvie D. Changing the environment to improve population health: a framework for considering exposure in natural experimental studies. J Epidemiol Community Health [Internet]. BMJ Publishing Group Ltd; 2016 Sep [cited 2016 Dec 9];70(9):941–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27056683
25. Ogilvie D, Bull F, Powell J, Cooper AR, Brand C, Mutrie N, et al. An Applied Ecological Framework for Evaluating Infrastructure to Promote Walking and Cycling: The iConnect Study. Am J Public Health [Internet]. 2011 Mar [cited 2016 Dec 9];101(3):473–81. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21233429
26. Silveirinha de Oliveira E, Aspinall P, Briggs A, Cummins S, Leyland AH, Mitchell R, et al. How effective is the Forestry Commission Scotland's woodland improvement programme--'Woods In and Around Towns' (WIAT)--at improving psychological well-being in deprived urban communities? A quasi-experimental study. BMJ Open [Internet]. British Medical Journal Publishing Group; 2013 Aug 30 [cited 2016 Dec 9];3(8):e003648. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23996826
27. 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.
28. 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 [Internet]. 2016

[cited 2016 Jul 27];16:35. Available from: http://dx.doi.org/10.1186/s12911-016-0274-7
29. Cornish RP, John A, Boyd A, Tilling K, Macleod J. Defining adolescent common mental disorders using electronic primary care data: a comparison with outcomes measured using the CIS-R. BMJ Open [Internet]. British Medical Journal Publishing Group; 2016 Dec 1 [cited 2016 Dec 9];6(12):e013167. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27909036
30. 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 [Internet]. Pergamon; 2017 May 1 [cited 2017 Nov 24];45:77–84. Available from: http://www.sciencedirect.com/science/article/pii/S135382921730223X?via%3Dihub
31. Lyons RA, Jones KH, John G, Brooks CJ, Verplancke J-P, Ford D V, et al. The SAIL databank: linking multiple health and social care datasets. BMC Med Inform Decis Mak [Internet]. 2009 Jan [cited 2014 Aug 19];9(1):3. Available from: http://www.biomedcentral.com/1472-6947/9/3
32. Ford D V, Jones KH, Verplancke JP, Lyons RA, John G, Brown G, et al. The SAIL Databank: building a national architecture for e-health research and evaluation. BMC Heal Serv Res. 2009;9(1):157.
33. Jones KH, Ford D V, Jones C, Dsilva R, Thompson S, Brooks CJ, et al. A case study of the Secure Anonymous Information Linkage (SAIL) Gateway: A privacy-protecting remote access system for health-related research and evaluation. J Biomed Inform [Internet]. 2014 Aug [cited 2014 Aug 19];50:196–204. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4139270&tool=pmcentrez&rendertype=abstract
34. Rodgers SE, Lyons RA, Dsilva R, Jones KH, Brooks CJ, Ford D V, et al. Residential Anonymous Linking Fields (RALFs): A Novel Information Infrastructure to Study the Interaction between the Environment and Individuals' Health. J Public Heal. 2009;31(4):582–8.
35. Ordnance Survey. AddressBase Premium | Business and government | Ordnance Survey [Internet]. 2014 [cited 2014 Mar 27]. Available from: http://www.ordnancesurvey.co.uk/business-and-government/products/addressbase-premium.html
36. Ordnance Survey. OS(R) MasterMap(R) Integrated Transportation NetworkTM Layer technical specification. (c) Crown copyright 2011 [Internet]. 2012. Available from: http://www.ordnancesurvey.co.uk/oswebsite/support/products/os-mastermap/itn-layer-technical-specification/index.html
37. 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 Heal Res [Internet]. 2016 Mar [cited 2016 Oct 27];4(3):1–184. Available from: http://www.journalslibrary.nihr.ac.uk/phr/volume-4/issue-3
38. Rodgers SE, Heaven M, Lacey A, Poortinga W, Dunstan FD, Jones KH, et al. Cohort profile: the housing regeneration and health study. Int J Epidemiol. 2012 Feb;43(1):52–60.
39. Macintyre S. Inequalities in health in Scotland: what are they and what can we do about them? MRC/CSO Soc Public Heal Sci Unit Occas Pap no 17, Glas. 2007;
40. Frank J. Public health interventions to reduce inequalities: what do we know works?

Can J Public Health. 2012 Sep;103(7 Suppl 1):eS5-7.
41. Mitchell R, Popham F. Effect of exposure to natural environment on health inequalities: an observational population study. Lancet [Internet]. 2008 Nov [cited 2016 Jul 28];372(9650):1655–60. Available from: http://linkinghub.elsevier.com/retrieve/pii/S014067360861689X
42. Welsh Government. Welsh Government | Welsh Index of Multiple Deprivation (WIMD) [Internet]. 2012. Available from: http://gov.wales/statistics-and-research/welsh-index-multiple-deprivation/?lang=en
43. Natural Environment Framework: Woodland case study - SERG_Welsh_woodland_creation_case_study.pdf.
44. Natural Resources Wales / Open Access Land.
45. Welsh Government | Improving opportunities to access the outdoors for responsible recreation.
46. West Rhyl Greenspace Project. Health Impact Assessment final report. 2014.
47. Gan M, Deng J, Zheng X, Hong Y, Wang K. Monitoring Urban Greenness Dynamics Using Multiple Endmember Spectral Mixture Analysis. PLoS One. 2014 Nov;9(11).
48. Beyer K, Kaltenbach A, Szabo A, Bogar S, Nieto F, Malecki K. Exposure to Neighborhood Green Space and Mental Health: Evidence from the Survey of the Health of Wisconsin. Int J Environ Res Public Health [Internet]. 2014 Mar 21 [cited 2016 Dec 9];11(3):3453–72. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24662966
49. Welsh Government | Technical Advice Note 16: Sport, Recreation and Open Space (2009).
50. Ordnance Survey. OS MasterMap Topography Layer [Internet]. 2017 [cited 2017 May 26]. Available from: https://www.ordnancesurvey.co.uk/business-and-government/products/topography-layer.html
51. OS Greenspace | OS GetOutside [Internet]. [cited 2017 Nov 24]. Available from: https://www.ordnancesurvey.co.uk/getoutside/greenspaces/
52. Rodgers SE, Heaven M, Lacey A, Poortinga W, Dunstan FD, Jones KH, et al. Cohort profile: the housing regeneration and health study. Int J Epidemiol [Internet]. 2014 Feb [cited 2014 Sep 10];43(1):52–60. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23179304
53. Chow S-C, Wang H, Shao J. Sample Size Calculations in Clinical Research. Second. CRC Press; 2007. 480 p.
54. Penn R, Berridge D. Social Statistics. Vol. 1. SAGE Publications, Inc; 2010. 642 p.
55. Poses RM, Smith WR, McClish DK, Anthony M. Controlling for confounding by indication for treatment. {Are} administrative data equivalent to clinical data? Med Care. 1995 Apr;33(4 Suppl):AS36--46.
56. Davies RB, Crouchley R. Control for Omitted Variables in the Analysis of Panel and Other Longitudinal Data. Geogr Anal. 1985 Jan;17(1):1–15.
57. Brunsdon C, Fotheringham AS, Charlton ME. Geographically weighted regression-modelling spatial non-stationarity. J R Stat Soc Ser D (The Stat. 1998;47(3):431–43.
58. Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, et al. Compendium of physical activities: an update of activity codes and MET intensities. Med Sci Sports Exerc. 2000 Sep;32(9 Suppl):S498-504.
59. Dugdill L, Stratton S. Evaluating Sport and Physical Activity Interventions. In: Physical Activity and Health Promotion: Evidence Based Approaches to Practice. Blackwell Publishing;

60. 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 (Baltim) [Internet]. 2016 Oct [cited 2016 Dec 9];91:383–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27658650