NIHR Public Health Research programme

NHS National Institute for Health Research

PHR Protocol - project ref: 09/3007/02
Version: 1
Date: November 2011

Change in alcohol outlet density and alcohol-related harm to population health

Chief investigator	Professor David Fone
Sponsor	Cardiff University
Funder	NIHR Public Health Research Programme
NIHR Portfolio number	
ISRCTN registration (if applicable)	N/A

Change in alcohol outlet density and alcohol-related harm to population health

1. Aims/Objectives:

Primary research question:

We propose to directly address the NIHR commissioning brief PHR 09/3007 research question "*What is the impact of a reduction in the availability of alcohol on community alcohol health related harm and/or consumption in the local community?"*

We will do this by investigating the effect of change in alcohol outlet density on important alcohol-related health outcomes in Wales.

The outcomes we propose to investigate are: levels of daily consumption, including above sensible limits and binge drinking; hospital admissions; accident & emergency (A&E) department attendances; and police data on violent assaults.

Secondary research questions:

Here we will directly address the NIHR commissioning brief to consider the impact on health inequalities. We suggest this has two important methodological components.

First, it is clear from published research that selective population migration may account for a spurious widening of health inequalities over time.[1,2] We will estimate the effect of population migration on the estimates of outlet density and hence on the associations with alcohol-related harm.

Second, we will assess the impact of change in outlet density on alcohol-related health inequalities, taking the effect of population migration into account.

Our secondary research questions are therefore:

1. Does a health selection effect from population migration at small-area level explain any observed associations between outlet density and alcohol-related harm?
2. What effect does change in outlet density have on population inequalities in alcohol-related health?

Objectives:

1. To update the review of the literature
2. To collate data on alcohol outlets in Wales for all 22 local authorities
3. To estimate outlet density using a network-based spatial analysis
4. To clean and validate data from the Welsh Health Surveys 2005-09
5. To prepare the datasets of alcohol-related hospital admissions, A&E department attendances, and police recorded violent assault
6. To carry out descriptive and multilevel regression analyses
7. To carry out spatial analyses
8. Final update of the literature review
9. To write reports/papers/abstracts/dissemination

2. Background and contribution of existing research:

Research investigating cross-sectional associations between higher outlet density and alcohol-related harm is well established. Less certain is the association between *change* in outlet density and alcohol-related harm. We carried out an initial scoping search, with a basic search strategy and limited number of information sources and search terms, of relevant research to focus our research question, aim, objectives, design and methods. We included research studies and reviews of research studies from the earliest date available in the database in developed (OECD) countries only.

Summary of findings
From a brief examination of titles and abstracts, 143 research publications were of broad relevance to the topic area, of which 78 looked directly at the effect of alcohol outlet density on health outcomes.

Of these 78 publications, 12 were reviews of studies and 66 were research papers (circa 62 individual studies) of which 55 (83%) were from the USA, 4 (6%) from Australia, and one each from New Zealand and Norway; the remaining five could not be assigned to a location but were assumed to be from developed countries. Eleven papers were not community based. No UK studies were found in this brief scoping search. Only one review article[3] and ten primary research papers[4-13] investigated *change* in outlet density and these are most relevant to the NIHR brief and our proposal.

Change in outlet density
The review article included papers published up to 2006 and summarised the effects of a change in outlet density resulting from three groups of studies: interrupted time series; natural experiments from alcohol bans, and changes in licensing arrangements.[3] Many of these studies were old and few accounted for spatial autocorrelation in their analyses (i.e. the lack of independence of proximate small-areas). This review did not include any methodological assessment of the measurement of outlet density. The overall conclusion was that an increase in outlet density led to increased consumption and inter-personal violence, but the evidence was less clear cut for an association with motor vehicle crashes.[3]

We found ten other papers investigating associations between change in outlet density and a limited range of outcomes, namely violence,[5,7,8,9,11,12,13] sexually transmitted infection,[4,10] suicide,[6] and car crashes.[7,11] Three of these papers[5,9,12] were included in a systematic review of all study designs.[14]

A six-year longitudinal study of change in outlet density and suicide set in California found that change in bar and off-premise outlet densities were positively related, and restaurant densities negatively related, to attempted and completed suicide rates.[6] In this study outlet density was measured simply as the number of outlets/per zip code population, with no spatial analysis. Time-lagged random effects models correcting for spatial autocorrelation (i.e. where variables in adjacent areas tend to be correlated) were used.

Using the same dataset a second paper reported that change in each measure of outlet density was positively related to likely alcohol-related accident rates, using Police-recorded crashes judged to include alcohol, and hospital admissions resulting from motor vehicle crashes identified from ICD-10 codes E810-E825 (which are not alcohol-specific).[11] A smaller study of alcohol-related crashes two years before and after closure of drive-up alcohol outlets in New Mexico reported a non-significant association.[7]

A third analysis of the same dataset found that increases in bar and off-premise outlet densities were positively related with an increase in the rate of violence, defined as a

hospital admission with an overnight stay with ICD-10 codes E960-969 for injuries resulting from inter-personal violence, which again are not alcohol-specific.[5]

A study set in Melbourne found that an increase in the number of outlets over a nine-year period was positively related to an increase in numbers of violent assaults recorded by the Police as taking place between 8pm and 6am (used as a proxy for alcohol-related assault for which specific data were not recorded).[8] The fixed effects analyses, which controlled for spatial autocorrelation, modelled numbers of assaults and outlets rather than rates or density. This analysis further categorised postcodes into five clusters, based on socio-demographic factors, to investigate more detailed relationships and found the strongest associations in central and inner-city suburbs. An earlier study from Norway analysing data from 1960 to 1995 found in a time-series analysis that outlet density was significantly associated with violent crime.[9]

Finally, research from Los Angeles reported that a reduction in alcohol availability within census tracts (resulting from a riot in 1992) was associated with a significant reduction in gonococcal infection rates,[4] partially mediated by neighbourhood social capital.[10] A significant reduction was also found for violent assaults[12] but a technical statistical paper published one year later suggested that the association reported in their earlier analysis of violent assaults was no longer significant after using a different modelling method.[13]

Measurement of outlet density
In the longitudinal studies, outlet density was estimated simply as the number of outlets per resident population in seven of the eight studies which estimated a density.[5-7,9-11,13] Other studies estimated density using the number of road-miles[4,12] or the geographical area[12] as the denominator. One study, which also used a population denominator, reported no difference between methods.[12]

Type of spatial model
None of the ten papers investigating associations between change in outlet density and alcohol-related outcomes used any type of spatial network-based measure of outlet density or conducted a spatial analysis.

3. Need:

The NIHR commissioning brief research question has not been answered in published research. We did not identify any UK studies in our scoping review. Although there is cross-sectional evidence to suggest that higher outlet density is associated with alcohol-related harm, the evidence from longitudinal studies for associations between harm and a change in outlet density is less well established.

There are methodological problems inherent in this type of study. Consideration must be given to the measurement of outlet density. If the data permit, the numerator should distinguish between different types of premise and the choice of density measure ranges from a simple population count, geographical area or road miles in an administrative area to a spatial measure of exposure in which the number of outlets are summed within a buffer zone defined by distance, walking or drive time for each study subject. New measures of outlet density are required that incorporate spatial accessibility. Statistical modelling should allow for the effects of spatial autocorrelation and spatial models are required to investigate spatial effects.

4. Methods:

a. Setting

The Setting is Wales, with 22 local authorities, total population 3 million; population 2.4 million aged 16 years and over (2001 Census).

b. Design

The intervention we propose to evaluate is the change in the availability of alcohol, measured as outlet density. The commissioning brief specifies the intervention should consider a "reduction in availability of alcohol..." but we propose to extend this to a two-tailed situation where change can measure either a reduction or an increase in alcohol availability. We describe in detail in section 4d how we propose to estimate outlet density in this project.

The complex intervention takes the form of a natural experiment as is common in public health contexts, for example in changes to policy on the availability of alcohol. These are not amenable to randomisation but have a natural non-random variability in geographical location and time. Different local authorities have different policies and we will obtain a comprehensive list of these policies and their variations, for example a specific policy for handling applications for licences in areas where there is a danger of saturation. These variations in local policy will be associated with increases and decreases in outlet densities and we plan to contrast the effects of different local policies in our exposure measure and hence in the health outcomes.

We propose four studies**:**

Study 1: Primary outcome: alcohol consumption

A longitudinal analysis of change in levels of alcohol consumption using data from the 75,000 people aged 16 years and over who anonymously responded to the five annual Welsh Health Surveys carried out between 2005 and 2009.[15-19] Consumption will be categorised as an ordinal scale of no alcohol, below sensible limit, above limit, and binge drinking using the Department of Health definition 2007.[20] (see section 4c).

Study 2 Secondary outcome: hospital admissions

A longitudinal analysis of five years of hospital in-patient data for residents of Wales aged 16 years and over using the Patient Episode Database for Wales (PEDW), 2005-09. This includes hospital admissions for alcohol-related causes (around 13,000/year),[21] defined by a published set of ICD-10 codes.[22]

Study 3 Secondary outcome: Accident & Emergency attendances

A longitudinal analysis of residents of Wales aged 16 years and over attending an A&E department at night (10pm – 6am - around 100,000/year) as a proxy for alcohol-related injury between 2009 and 2011.

Study 4 Secondary outcome: alcohol-related violent crime

A longitudinal analysis of five years of Police data on alcohol-related violent crime against the person in Wales between 2005 and 2009 (around 47,000/year).

c. Data collection

Use of existing record-linked datasets: the Secure Anonymous Information Linkage Databank

The Secure Anonymous Information Linkage (SAIL) databank held and managed within the Health Information Research Unit (HIRU) at Swansea University contains health, social and education data on three million residents of Wales and currently includes thirteen datasets containing nearly one billion records.[23,24] The smallest geographical area for which data are already linked and may be released is the 2001 Census Lower Super Output Area (LSOA).[25] Point data, when in combination with health data in the SAIL databank, cannot be anonymised for release. Confidentiality issues are discussed in section 6.

Welsh Demographic Service (WDS)

The WDS 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 update of the address record are held, thereby providing durations of residency for several different homes and the ability to link to local environment exposures at each. Demographic data may be used to create population sub-groups based on age, gender and location for the required date or duration. The WDS contains address information linked anonymously at the individual level (the anonymised linking field, 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. [23,24]

Residential Anonymous Linking Fields

A particular strength of the SAIL system is the development of the system 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 Linking Fields (ALF) and Residential Anonymous Linking Fields (RALF). Address data are matched at NWIS where identifiable addresses are replaced with RALFs.[26] The residence-based metrics are then fully incorporated into the SAIL databank by linking RALFs to ALFs.

Environment Geographic Information System

In parallel, but separate from the SAIL databank, is the Environment Geographic Information System (eGIS) which contains map data from the Ordnance Survey at high spatial resolution. Ordnance Survey Master Map (OSMM) Address Layer 2[27] contains point data for all residences and the OSMM Integrated Transportation Network dataset.[28] The Address Layer 2 point data layer contains a point for each residence, which is placed within the footprint of the residence. The buildings are surveyed with a spatial accuracy of ± 1-2 m. This provides georeferences at a high spatial resolution.

Alcohol outlets

There are 22 local authorities in Wales who are responsible under The Licensing Act 2003 for maintaining public registers of Premises and Personal Licence holders, available annually from 2005. Under The Act these data are available for public inspection in council offices. Each authority keeps a database (electronic or hard-copy) of the name and address of each premise with the full unit postcode.

To prepare for this study, and confirm the feasibility of obtaining outlet information, we have approached the responsible officer in each of the 22 local authorities in Wales asking to share the data on outlets. All 22 local authorities have agreed to supply the data. We have already received electronic data from eight authorities for the study period. Of these, data from four authorities are in a format that can be analysed with little further cleaning. Only two authorities distinguished between on and off-sales and so we will not have this categorisation for all-Wales data.

Using the full unit postcode we can georeference each outlet to within any geographical boundary using a standard geographical information system (GIS). These postcodes have been validated by the local authorities and we will check each postcode to the address using the OSMM Address Layer 2.[27] In this study the basic geographical unit for small-area analyses will be the 2001 Census LSOA. These areas vary in size but contain an average of 1500 people, with the constraint of a minimum population of 1000.[25] There are 1896 LSOAs in Wales.

To assess the magnitude of change in outlets over time we have assessed the outlet data on licensed and off-sales premises which have already been supplied by four local authorities, with a total of 273 LSOAs (mean 13 outlets). The mean five-year change for local authorities was 6% (range 5% to 46%) which exceeds the 4% change in outlets over a six-year period reported in the study of change in outlet density which gave sufficient power to detect change in a rare (suicide) outcome[6] and the -0.4% mean change in outlets over a six-year period reported in a second longitudinal study.[11]

The data received from the other four authorities are fully postcoded but need converting from text files and duplicate entries removed. Ten further authorities have confirmed that their data will be made available in electronic format. Four authorities have confirmed that the data will be available only in hard copy and so we will scan these data (approximately 2,000 A4 pages) using Optical Character Recognition software. Our pilot investigations and discussions with the lead person in each authority have shown that the systematic collation of the alcohol outlet data is feasible.

Primary outcome: Welsh Health Survey
Data on consumption will be used from the Welsh Health Surveys 2005-09. We are very familiar with these data, having validated the geographical hierarchy in the 2005-2008 dataset and we have published analyses using Welsh Health Survey data.[29,30]

Participants were asked to state the highest number of units they had drunk on any one day in the previous seven days, using a standard prompt to convert different types and quantities of alcoholic drinks into units. The classification of units into ordinal categories of consumption used in the Welsh Health Survey dataset is based on the Department of Health definitions:[20]

Category	*Maximum units drunk on any day in the last week*
None	Did not drink in the last seven days
Sensible	Drank something; men drinking no more than 4 units, women no more than 3 units
Above guidelines, less than binge	Men drinking more than 4 and up to and including 8 units, women more than 3 and up to and including 6 units
Binge	Men drinking more than 8 units, women more than 6 units

Secondary outcomes:

Hospital admissions
The SAIL databank includes the complete Patient Episode Database for Wales (PEDW), which is the database of all inpatient and day case admissions for residents of Wales. The pseudonymisation process results in a unique, but unidentifiable, ID that enables a patient-based analysis rather than an admissions-based analysis.

Each hospital spell has an attached ICD-10 diagnosis code in one of 14 positions: primary, subsidiary, and 3rd to 14th position. The Office for National Statistics has published a list of ICD-10 codes to define causes of death which are exclusively alcohol-related i.e. not just where alcohol could be a contributory aetiological factor.[22] This definition has previously been used in Wales to quantify the annual number of alcohol-related admissions.[21]

There are several other published classifications of 'exclusively' alcohol-related conditions with considerable overlap between them. The differences from the ONS definition largely relate to non-fatal conditions excluded from the ONS cause of death list. In addition are the external cause codes for motor vehicle crashes, E810-E825, and injuries resulting from inter-personal violence E960-969. However these codes are not alcohol-specific. In the analysis we will define alcohol-related as the combined causes of death and condition codes and we will perform secondary analyses including ICD-10 'E' codes for the motor vehicle crashes and injuries resulting from inter-personal violence.

Each PEDW record is linked to the LSOA of residence to which a deprivation score, such as the Welsh Index of Multiple Deprivation 2005,[31] has been attached in SAIL. We can then attribute admissions to an LSOA and deprivation fifth for the analysis.

Accident & Emergency department data
The Accident & Emergency Department Data Set (EDDS) is a new dataset within HIRU/SAIL and the first wave of data for the 13 Accident & Emergency (A&E) Departments in Wales for 2009 has now been cleaned and validated. Data for subsequent years will be entered into SAIL as it is received, currently on a monthly basis. The data record attendances for all injuries with date and time of day. There is no definitive uniformly applied code for alcohol-related attendances in this dataset and so we will use attendance at night (10pm to 6am) as a proxy for alcohol-related as the majority of injury attendances at night (particularly aged 16-39) are likely to be alcohol related. This assumption has been used in previous research,[5] and is the best possible with the data available. These attendances will be linked to RALFs and hence to anonymised network measures for the analysis. Using ALFs/RALFs we will also trace attendees who are subsequently admitted to hospital (by linking to PEDW) and exclude all those who do not have an alcohol-related ICD-10 code.

The numbers are sufficient for a spatial analysis: there were 101,908 night time attendances at the 13 A&E departments in Wales in the financial year 2009/10. The highest numbers were on Fri/Sat and Sat/Sun nights (15,862, 17,098) compared to weekdays (weekday range 13,374 to 14,232), weekly mean 14,558, SD 1386).

Police data
Police forces record electronically all incidents reported to them. Incidents of violence against the person are detailed according to incident type (e.g. assault occasioning actual bodily harm), the injury incurred, incident location and alcohol use. These incidents are recorded at the premises, address or street location level. Eastings and Northings locations are also available, although the accuracy of these coordinates has not been evaluated. Our work with police data on an exploratory trial of a disorder reduction intervention found the data to be sufficiently robust for research purposes.[32] Data sharing agreements between Cardiff University and South Wales Police are currently in place and will be extended to the other three police forces in Wales. Although a proportion of violent incidents are not reported to the police, these data remain the most detailed and accurate records of violent incidents at the premises or street level.

Pilot work using South Wales Police data was conducted to establish the feasibility of using police recorded incident data for the period of the proposed study, 2005-09. Approximately 42% of incidents were recorded as involving alcohol on the part of the offender or the victim. While this rate is slightly below the 50% prevalence among incidents recorded by the British

Crime Survey[33] we believe that these data are of sufficient quality for our study. Since 99% of incidents are described at the street level and 100% of incidents are described at the police beat level we can map and aggregate the data at LSOA-level. Furthermore, the recording of incident type and presence of injury allows the study to explore the role of outlet density for both occurrence and severity.

The Police violent crime data at LSOA-level will be added to SAIL for this project.

d. Data analysis

Spatial modelling of accessibility of the population to alcohol outlets to estimate outlet density

Our scoping literature review has found four main approaches to estimate outlet density: (1) number of outlets in an administrative per capita or per unit geographical area, (2) number of outlets per network miles per geographical area, (3) number of outlets per geographical area defined by walking or drive time (a 'buffer zone'), and (4) closest Euclidian distance from each subject's home to an outlet.

We wish to estimate a new measure of outlet density which includes a more realistic and robust measure of accessibility than previously published methods. We will estimate a summary statistic for distance to all outlets from each residence within a defined buffer zone. The measures of distance calculated will be network distances and travel times which will more accurately measure exposure than Euclidian distances.

We therefore propose a spatial methodology to estimate outlet density using the unique advantages of the WDS population data within the SAIL databank:

(1) Using the service area tool in ArcGIS Network Analyst module, together with OSMM data and point location facility data in the eGIS[27,28] we will calculate the mean network distance for each individual to each outlet within a meaningful network range (10 minutes walk and 10 minutes drive-time) of each residence (RALF). The individual-level measure of outlet density exposure will be a function (e.g. reciprocal) of the mean distance

(2) The LSOA density can be derived by averaging the values for all individual residents in the LSOA

The primary analysis will be for all alcohol outlets but we will also estimate outlet densities for on and off-sales separately for those authorities who can supply this categorisation.

Proposed sample size

As noted, outlet data received to date from four local authorities showed a mean five-year change of 6% (range 5% to 46%) which exceeds the % change in outlets reported in previous studies to have sufficient power to detect change in rare outcomes.[6,11] Thus we will have sufficient change in outlet density for the analysis and to answer the research questions.

For the primary outcome, based on data from five Welsh Health Surveys, the sample size is approximately 75,000. Because we do not yet have the data on the location of alcohol outlets by year, we do not know the distribution of changes and so cannot complete a conventional sample size calculation. However we previously carried out an analysis of similar data of ca. 60,000 respondents to four successive Welsh Health Surveys 2005-2008 and can draw on that experience. In similar multilevel models to those anticipated here, for the primary outcome analysis, standard errors were typically between 0.04 and 0.06, varying between model parameters. These led to narrow confidence intervals for parameter values; for example the estimated odds ratio associated with a quintile of deprivation was 0.82 with 95% confidence interval (0.76, 0.88). So with even more data here the precision will be

greater and even small differences in outcomes will be recognised.
For the other outcomes the study population is essentially the 2.4 million people living in Wales and aged at least 16 years. The outcomes are not rare with approximately 13,000 alcohol-related admissions, 100,000 night time A&E attendances, and around 47,000 violent crimes against the person in Wales annually. We will have very high power for detecting small differences between small-areas (LSOAs) at the extremes of changes in alcohol outlet density. For example there would be over 80% power for identifying a reduction from 5.4 per thousand per annum to 5 per 1000 per annum between the extreme quintiles of change in outlet density. While this ignores clustering within LSOA, that effect is unlikely to be so large that the power is greatly reduced and there will be more power for detecting a larger difference. If the correction factor for clustering was two, a common finding in work of this type, the required sample size would effectively be doubled. However with the planned sample size there would still be over 80% power for detecting a reduction from 5.4 per thousand per annum to 4.8 per thousand per annum in the admission rate for alcohol-related conditions.

Statistical analysis
The primary analysis will be for all alcohol outlets but we will also analyse the data using outlet densities for on- and off-sales separately for those authorities who can supply this categorisation. Our analysis plan has to take into account the nature of the data available through SAIL. Data on individuals nested within households nested within LSOAs will be available, but the geographical details will be anonymised, so that the geographical location will not be known. This has all of the advantages of individual outcome and exposure data, together with demographic details, and has not been performed previously, to the best of our knowledge. It has the disadvantage, however, that a full spatial analysis will not be possible. For a spatial analysis we will have to follow other authors and use an ecological approach using data aggregated over LSOAs. Our analysis plan is:

(1) Full descriptive statistics

(2) A spatial model based on LSOA-level data to explore the relationship between counts/rates of the four outcome measures and outlet density. The models will account for the year and will adjust for aggregated population characteristics, such as the percentages of males and of certain socio-economic groups, and the median age. A variety of spatial models can be used for such data to allow for spatial autocorrelation, including the Bayesian approach of Besag, York and Mollié,[34] the random effects models developed by Gruenewald and colleagues and used in the three Californian longitudinal studies[5,6,11] and geographically weighted regression (GWR) models.[35-37] GWR is slightly different from other spatial regression techniques such as the spatially lagged model and spatial error model[38] 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 the outlets in explaining the spatial distribution of health outcomes. The process of GWR modelling will be implemented within the eGIS environment, using ArcGIS 9.3 extension of the GWR software. The 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.

(3) A multilevel analysis using the hierarchy outlined above of individuals nested within households nested within LSOAs, nested within local authorities. We will use multilevel ordinal models of the defined categories of consumption for the WHS data and multilevel logistic models of the odds of hospital admission and A&E attendance as a function of change in individual outlet density exposure, adjusting for confounding individual variables (age, gender, SES) and LSOA deprivation score. This method cannot be used for the crime data as we only know location of offence with no information on individuals. Multilevel modelling is a well-established technique, see for example Goldstein 2003,[39] and we will fit

models using the standard MLWiN[40] and R software.[41]

In order to address the secondary research questions:
(4) We will assess the impact on health inequalities by firstly stratifying the models in (2) and (3) above by Welsh Index of Multiple Deprivation, using quintiles of deprivation to examine whether the modelled associations vary with deprivation category. This method was used for postcode sectors in the Australian longitudinal analysis.[8] A better method which we will use is to model the interaction between LSOA deprivation score and change in outlet density and the two-way interaction including individual SES (using the NS-SEC).[42] This method uses the whole dataset and the coefficient of the interaction terms gives information on the effect on inequalities over time.

(5) Since the annual estimation of outlet density exposure automatically includes people who move residence, (from using the WDS within SAIL), our main models will correct for population migration. We will assess the impact of population migration by re-estimating annual outlet densities using the baseline population, repeating the models above, and comparing to the main models. Health selection bias can thus be assessed.

Presentation and mapping of the results
An important element of the research will be mapping of the data. This will form part of the substantive investigation as well as part of the presentation and communication of the results. It is well established that mapping health and related social and economic data can provide important insights into the causal processes that affect health outcomes.[43] Global and local statistical measures of spatial clusters can be misleading without mapping,[44,45] which provides an excellent method of revealing potential anomalies in the data. The research will therefore use mapping techniques in the specification and testing of the alcohol outlet density variable and the spatial model.

Mapping will also be used as part of the diagnostic tests and interpretation of the multilevel models, principally through mapping the LSOA residuals. This is an important element of model building[46] and can be used to refine the specification of the model as well as identify structure not accounted for by the explanatory variables. For instance, systematic variation in the size and direction of residuals across space can be indicative of spatial autocorrelation and/or spatial heterogeneity caused by mis-specification of the model. Mapping the residuals and analysing them using spatial statistical techniques such as local indicators of spatial autocorrelation can prove invaluable in model building. Moreover, mapping will allow a visual analysis of the spatial correlation between the residuals and the location of the alcohol outlets. A clear relation, such as concentrations of large positive/negative residuals around clusters of alcohol outlets, could indicate social processes unexplained by the model and lead to further analysis.

Maps will also form an important part of the presentation and communication of the results to different audiences. Politicians and policy makers respond well to visual information and maps have the capacity of presenting a vast amount of complex information in an efficient manner. This is true of health data where the spatial correlations between alcohol outlets and health outcomes can be visually demonstrated. Maps will help explain our methodology and confirm the findings of our models.

	2011						2012												2013												2014
Month Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**
Month Activity	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D	J
1.Project team formal research meetings																															
2.Study steering group meetings																															
3.Data access agreement to use the WHS																															
4.Prepare WHS complete dataset for analysis, validation of geography																															
5.Update literature review																															
6.Update WDS data in SAIL databank, preparation of dataset of alcohol-related hospital admissions, SQL extraction																															
7.Preparation of dataset of A&E attendances, SQL extraction																															
8.Preparation of dataset of violent assaults																															
9.Collation and preparation of outlet data from 22 local authorities																															
10.Develop the network-based spatial methods for outlet density																															
11.Descriptive analyses																															
12.Multilevel regression analyses																															
13.Spatial analyses																															
14.Final literature review update																															
15.Write reports/papers/abstracts																															
16.Submit report to NIHR																															
17.Start dissemination of results																															

6. Project Management:

Research Governance and ethical arrangements
The information governance requirements for use of the Secure Anonymised Information Linkage (SAIL) dataset in this proposal have been approved by the Independent Information Governance Review Panel (IGRP).

Participant consent is not required because all the SAIL study outcome data (on hospital admissions, A&E attendances and Police data) will be anonymised before they are incorporated into the SAIL databank and anonymously linked with other SAIL datasets. As the SAIL databank is fully anonymised, it does not fall into the remit of the National Information Governance Board who provide section 251 (formerly section 60) exemption to use identifiable data without consent.

The anonymised Welsh Health Survey dataset is supplied under a formal Data Access Agreement between Cardiff University and the Welsh Government. Welsh Health Survey data do not have permission from the Data Owner (Welsh Government) to link into SAIL at individual-level but can be linked, using data aggregated by LSOA, to outlet density exposures.

The Chairman of the Research Ethics Committee for Wales has confirmed in writing that this study does not require ethical approval under NHS research governance arrangements.

Study steering committee
A multi-disciplinary and multi-agency study steering committee (SSC) will be established to oversee and guide project activities. The SSC will be constituted to include the lead and one co-applicant from each School in Swansea and Cardiff University, and a member from the NIHR, the study sponsor (Cardiff University), South Wales Police, Cardiff Community Safety Partnership, the Involving People network (see section 7), and experts in the field of alcohol and violence research.

7. Service users/public involvement:

Public involvement is an important part of the project. The study outcomes which define patients – A&E attendances or admission to hospital – do not include recruited patients who have consented to participate but rather we have permission for an anonymised secondary analysis of record-linked datasets. The same applies to victims of violent crime. The scope for patient involvement is therefore limited. However we plan to offer the charity Alcohol Concern the opportunity to provide patient perspectives.

The primary vehicle in Wales for public engagement, with whom we will work closely, is The Clinical Research Collaboration-Cymru "Involving People" network (http://www.involvingpeople.org.uk/). Involving People aims to provide input into the strategy, development and implementation of health and social care research in Wales.

8. References:

[1] Norman P, Boyle P, Rees P. Selective migration, health and deprivation: a longitudinal analysis. *Social Science & Medicine* 2005;60(12):2755-71.
[2] Connolly S, O'Reilly D, Rosato M. Increasing inequalities in health: is it an artefact caused by the selective movement of people?. *Social Science & Medicine* 2007;64(10):2008-15.
[3] Campbell CA, Hahn RA, Elder R, et al. Guide to Community Preventive Services: The Effectiveness of Limiting Alcohol Outlet Density as a Means of Reducing Excessive Alcohol Consumption and Alcohol-Related Harms. *American Journal of Preventive Medicine* 2009;6:556-69.
[4] Cohen DA, Ghosh-Dastidar B, Scribner R, et al. Alcohol outlets, gonorrhea, and the Los Angeles civil unrest: A longitudinal analysis. *Social Science & Medicine* 2006;62(12):3062-71.
[5] Gruenewald PJ, Remer L. Changes in outlet densities affect violence rates. *Alcoholism: Clinical and Experimental Research* 2006;30(7):1184-93.
[6] Johnson FW, Gruenewald PJ, Remer LG. Report of results from the OSHPS reabstracting project: an evaluation of the reliability of selected patient discharge data July through December 1988. *Alcoholism: Clinical and Experimental Research* 2009;33(12):2124-33.
[7] Lapham SC, Gruenwald PJ, Remer L, Layne L. New Mexico's 1998 drive-up liquor window closure: Effect on alcohol-involved crashes. *Addiction* 2004;99(5):598-606.
[8] Livingston M. A longitudinal analysis of alcohol outlet density and assault. *Alcoholism: Clinical and Experimental Research* 2008;32(6):1074-9.
[9] Norstrom T. Outlet density and criminal violence in Norway, 1960-1995. *Journal of Studies on Alcohol* 2000;61(6):907-11.
[10] Theall KP, Scribner R, Ghosh-Dastidar B, Cohen D, Mason K, Simonsen N. Neighbourhood alcohol availability and gonorrhea rates: Impact of social capital. *Geospatial Health* 2009;3(2):241-55.
[11] Treno AJ, Johnson FW, Remer LG, Gruenewald PJ. The impact of outlet densities on alcohol-related crashes: A spatial panel approach. *Accident Analysis and Prevention* 2007;39(5):894-901.
[12] Yu QZ, Scribner R, Carlin B, et al. Multilevel spatio-temporal dual changepoint models for relating alcohol outlet destruction and changes in neighbourhood rates of assaultive violence. *Geospatial Health* 2008;2(2):161-72.
[13] Yu QZ, Li B, Scribner RA. Hierarchical additive modeling of nonlinear association with spatial correlations-An application to relate alcohol outlet density and neighborhood assault rates. *Statistics in Medicine* 2009;28(14):1896-912.
[14] Popova S, Giesbrecht N, Bekmuradov D, Patra J. Hours and days of sale and density of alcohol outlets: Impacts on alcohol consumption and damage: A systematic review. *Alcohol and Alcoholism* 2009;44(5):500-16.
[15] National Assembly for Wales. *Welsh Health Survey 2004/05*. National Statistics, 2006.
[16] Welsh Assembly Government. *Welsh Health Survey 2005/06*. National Statistics, 2007.
[17] Welsh Assembly Government. *Welsh Health Survey 2007*. National Statistics, 2008.
[18] Welsh Assembly Government. *Welsh Health Survey 2008*. National Statistics, 2009.
[19] Welsh Assembly Government. *Welsh Health Survey 2009: Initial Headline Results.* [http://wales.gov.uk/docs/statistics/2010/100525sdr742010en.pdf]
[20] Department of Health, Home Office, Department for Education and Skills and Department for Culture, Media and Sport. *Safe. Sensible. Social. The next steps in the National Alcohol Strategy*. London: DH Publications, 2007.
[21] Gartner A, Cosh H, Gibbon R, Lester N. *A profile of alcohol and health in Wales.* National Public Health Service for Wales/Wales Centre for Health, 2009. ISBN: 0-9545544-9-3. [http://www.wales.nhs.uk/sites3/Documents/568/Alcohol%20and%20Health%20in%20Wales_WebFinal_E.pdf]
[22] Office for National Statistics. *Defining alcohol-related deaths*. ONS, 2006 [http://www.statistics.gov.uk/downloads/theme_health/Summary_responses.pdf]
[23] Ford D, Jones K, Verplancke J-P, Lyons R, John G, Brown G, et al. The SAIL Databank: building a national architecture for e-health research and evaluation. *BMC Health Services*

Research 2009;9(1):157.
[24] Lyons RA, Jones KH, John G, et al. The SAIL databank: linking multiple health and social care datasets. *BMC Medical Informatics and Decision Making* 2009;9(3):24.
[25] Office for National Statistics. *Census Geography. ONS,* 2007.
[26] Rodgers SE, Lyons RA, Dsilva R, et al. Residential Anonymous Linking Fields (RALFs): A Novel Information Infrastructure to Study the Interaction between the Environment and Individuals' Health. *Journal of of Public Health 2009*;10:1-7.
[27] Ordnance Survey. *OS MasterMap Address Layer 2: Technical specification*. 2007, p64.
[28] Ordnance Survey. *OS MasterMap Integrated Transportation Network Layer user guide and technical specification*. 2007, p99.
[29] Fone DL, Dunstan FD. Mental health, places and people: a multilevel analysis of economic inactivity and social deprivation. *Health & Place* 2006;12(3):332-44.
[30] Fone DL, Lloyd K, Dunstan FD. Measuring the neighbourhood using UK benefits data: a multilevel analysis of mental health status. *BMC Public Health* 2007,7:69.
[31] Local Government Data Unit - Wales*. Welsh Index of Multiple Deprivation 2005. Summary Report.* LGDU, 2005.
[http://new.wales.gov.uk/topics/statistics/theme/wimd2005/results/report-revised/?lang=en]
[32] Moore SC, Brennan I, Murphy S. Predicting and measuring premises-level harm in the night-time economy. *Alcohol Alcohol* 2011;46(3):357-63.
[33] Flatley J, Kershaw C, Smith K, Chaplin R, Moon, D. *Crime in England and Wales 2009/10*. London: Home Office, 2010.
[34] Besag J, York J, Mollié M. Bayesian image restoration with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* 1991;43:1-59.
[35] Brunsdon C, Fotheringham AS, Charlton ME. Geographically weighted regression: modelling spatial non- stationarity. *Journal of the Royal Statistical Society Series D-The Statistician* 1998;47(3):431-43.
[36] Fotheringham AS, Brunsdon C, Charlton ME. Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment and Planning A* 1998;30(11):1905-27.
[37] Fotheringham AS, Charlton ME, Brunsdon C. *Geographically Weighted Regression: the analysis of spatially varying relationships*. John Wiley, 2002.
[38] Ward MD, Gleditsch KS. Spatial Regression Models, Sage Publications, 2008.
[39] Goldstein H. *Multilevel Statistical Models* (3rd ed.). Edward Arnold: London, 2003.
[40] Rasbash J, Steele F, Browne WJ, Goldstein H. *A User's Guide to MLwiN, v2.10*. Centre for Multilevel Modelling, University of Bristol,2009.
[41] R Development Core Team: R: *A Language and Environment for Statistical Computing. R Foundation for Statistical Computing*. Vienna: Austria 2005, [http://www.R-project.org].
[42] Office for National Statistics. *The national statistics socio-economic classification user manual.* ONS, 2002. [http://www.statistics.gov.uk/methods_quality/ns_sec].
[43] Elliott P, Wakefield JC, Best NG, Briggs DJ. *Spatial Epidemiology: Methods and Applications*. Oxford University Press, Oxford, 2001.
[44] Anselin L. Local indicators of spatial association. *Geographical Analysis* 1995;27:93–116.
[45] Anselin L. "How (not) to Lie with Spatial Statistics" *American Journal of Preventive Medicine* 2006;30(2):S3-S6.
[46] Belsley DA, Kuh E, Welsch RE. *Regression diagnostics, identifying influential data and sources of collinearity*. Wiley: New Jersey, 2004.

This protocol refers to independent research commissioned by the National Institute for Health Research (NIHR). Any views and opinions expressed therein are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the PHR programme or the Department of Health.