Pilot study linking primary care records to Census, cardiovascular hospitalization and mortality data in Scotland: feasibility, utility and potential

Pilot study linking primary care records to Census, cardiovascular hospitalisation and mortality data in Scotland: feasibility, utility and potential

Anne Douglas (Research Fellow)[1], Genevieve Cézard (Research Fellow)[1], Colin R Simpson (Senior Lecturer)[1], Markus FC Steiner (Research Fellow, Clinical)[2], Raj Bhopal (Professor of Public Health)[1], Narinder Bansal (Senior Cardiovascular Scientist)[3], Aziz Sheikh (Professor of Primary Care Research & Development)[1], Hester JT Ward (Consultant in Public Health Medicine) [4], Colin M Fischbacher (Clinical Director for Information Services)[4], on behalf of the SHELS researchers

[1] Edinburgh Migration, Ethnicity and Health Research Group, Usher Institute of Population Health Sciences and Informatics, The University of Edinburgh, Teviot Place, Edinburgh EH8 9AG, UK

[2] Department of Child Health, School of Medicine, University of Aberdeen, Aberdeen, AB25 2ZG, UK

[3] Cardiovascular Epidemiology Unit, University of Cambridge, Cambridge, CB1 8RN, UK

[4] Information Services Division, NHS National Services Scotland, Edinburgh, EH12 9EB, UK

Corresponding author

Anne Douglas: anne.douglas@ed.ac.uk

This is a pre-copyedited, author-produced PDF of an article accepted for publication in Journal of Public Health *following peer review. The version of record* (*Pilot study linking primary care records to Census, cardiovascular hospitalization and mortality data in Scotland: feasibility, utility and potential*; Anne Douglas; Genevieve Cezard; Colin R. Simpson; Markus F.C. Steiner; Raj Bhopal; Narinder Bansal; Aziz Sheikh; Hester J.T. Ward; Colin M. Fischbacher; Journal of Public Health 2015; doi: 10.1093/pubmed/fdv192)

is available online at:

Published in Journal of Public Health 30/12/2015.

Abstract:

http://jpubhealth.oxfordjournals.org/cgi/content/abstract/fdv192?ijkey=2p0RDxafZzhwaIH&keytype=ref

Full Text:

http://jpubhealth.oxfordjournals.org/cgi/content/full/fdv192?ijkey=2p0RDxafZzhwaIH&keytype=ref

Abstract

Background

There are substantial ethnic variations in the risk of cardiovascular disease (CVD) related hospitalisation and mortality in Scotland. We piloted extracting and linking primary care risk factors to Scottish Census and health data, to test the feasibility of further investigating these variations.

Methods

Data extracted from 10 general practices were linked at individual level to Census and hospitalisation/death records. Linkage rates, reasons for non-linkage, and completeness of primary care data were examined. CVD risk ratios were calculated, adjusting for age, socio-economic status and primary care derived risk factors.

Results

Practice enrolment and data extraction proved challenging. Primary care records for 52,975 (55.2%) people were linked to Census data. Completeness and validity of risk variables were similar across ethnic groups. 48,325 (91.2%) of records had a valid smoking status recorded and 2,900 (5·5%) people had a primary care record of diabetes. Ethnic-specific adjusted estimates of CVD risk were plausible and consistent with previous work.

Conclusions

Risk factor data extracted from primary care were of good quality and successfully linked to national Census records. Given further methodological refinement, this method illustrates the potential value of linkage using national primary care datasets to contribute to public health surveillance and research.

Background

Studying inequalities in population health requires accessible, low cost, up-to-date data including health outcomes, risk factors and population characteristics. Quantifying and understanding such inequality helps identify population health needs and provides information for targeting services and research. Reducing health inequalities is a public health priority.[1] In relation to ethnic inequalities this requires accurate information on disease risks by ethnic group. There is growing evidence of ethnic inequalities in health globally[2-5] although still a shortage of data from Europe. The main limitations of previous research include using country of birth as a proxy for ethnic group – which is increasingly inaccurate in current multicultural populations – and lack of individual level socio-economic and risk factor data for use in risk estimates in different ethnic groups.[6,7]

The Scottish Health and Ethnicity Linkage study (SHELS) has linked National Health Service (NHS) Scotland hospital discharges and mortality to the 2001 Scottish Census, to explore ethnic variations in Scotland for priority health areas, including cardiovascular disease (CVD).[6,8-10] Access to Census data on social circumstances enabled our analyses to adjust for age and socio-economic status.[11] However, understanding the role of other risk factors in explaining these ethnic variations is vital. Some information, such as smoking, is not routinely available in hospital discharge data but is recorded in primary care. Research from the UK using primary care research databases[12-14] and internationally,[15,16] has shown the quality and utility of general practice data. There are current initiatives to create comprehensive national primary care datasets throughout the UK, and primary care data are therefore likely to become more widely available in the future.[17] We found no published studies linking primary care to national Census data (at individual level), or to other health datasets, to examine the relationship between ethnicity and health as the primary analysis.

The aims of this pilot study were: first, to assess the feasibility of accessing, extracting and linking primary care data to our existing 2001 Scottish Census and CVD hospitalisation and mortality datasets; and second, to examine the quality and completeness of the primary care risk factor data and assess the usefulness of these data in risk analysis models, interpreting the results in relation to previously shown ethnic variations in CVD and mortality outcomes.[6,8-10] We focused on Other White British and Pakistani groups, the largest minority populations in Scotland and, as exemplars, we examined the potential role of smoking and diabetes, two important risk factors for CVD.

Methods

The SHELS linkage methods have been published.[18,19] We followed a protocol that preserved anonymity, maintaining separation of personal identifiers from Census and clinical data. We used probabilistic linkage techniques matching names, addresses and dates of birth to link the 2001 Scottish Census to the Community Health Index (CHI), a register of patients using NHS Scotland. In 2008, CHI and Census numbers were encrypted and personal identifiers removed, leaving a look-up table. About 95% (approximately 4.65 million) of people participating in the Census were linked to CHI, with 85% or more linked in every ethnic group.

Between May 2011 and June 2012, 17 practices in NHS Lothian and NHS Greater Glasgow & Clyde Health Boards with above average proportions of ethnic minority populations were approached to participate in this project. The research team held telephone discussions with practices and an invitation letter, information sheet and practice authorisation form were then sent to those that expressed an interest in participating. Participating practices were offered a payment of £150 to contribute towards costs.

A trusted third party, Albasoft Ltd, extracted data from practices (Supplementary table 1). The data extract included CHI number but no patient or practice identifying details. After CHI numbers were replaced with an encrypted CHI, the resulting primary care dataset was linked to the 2001 Census records using a key look-up table at National Records Scotland (NRS) (Figure 1), creating a retrospective cohort study comprising 52,975 individuals. This was then linked to our pre-existing CVD hospitalisation/death datasets. [6,8-10] Encrypted CHI was removed from the data prior to any analysis.

Ethnicity, demographic and socio-economic factors

The ethnic group categories were those of the 2001 Census (Box 1) as reported by the head of household or individuals. We minimised aggregation to avoid missing important ethnic differences, but where numbers were small we combined Bangladeshis with Other South Asians and the Caribbean, African and Black Scottish or Other Black as 'African origin'. Age, sex and socio-demographic factors were available from Census data.

Smoking and diabetes data

Smoking status was defined for the linked cohort of 52,975 people using appropriate codes from the Read coding system[20] (Supplementary table 2) and the associated date. Status was classified as "ever

smoker" (at least one current smoker or only an ex-smoker Read code) or "never smoker" (only a never smoker Read code) between May 2001 and April 2010. CVD outcome data were only available up to 2008, so diabetes status was assigned using the last relevant Read code (Supplementary table 3) before the CVD event date. For those with no CVD event, the presence of any diabetes Read code during 2001-08 was used to assign diabetes status.

Statistical analysis

We calculated age-adjusted lifetime prevalence of diabetes and prevalence rate ratios by ethnicity and sex using Poisson regression with robust variance. We identified first CVD hospital admission or death over the seven year period from May 2001 to April 2008. First events were defined as those with no similar events in the preceding 10 years. Risk ratios (RR) for first CVD events were calculated, by sex, adjusting for age and education, and then for smoking and diabetes status. Risk ratios were multiplied by 100 (for easier interpretation), so that the White Scottish population, the reference group, had a RR of 100. Analysis was restricted to those aged 30 years and over. We used the statistical package SAS V.9.3 (SAS Institute Inc, Cary, North Carolina, USA).

Ethics, approvals and statistical disclosure

The work was approved by the Multicentre Research Ethics Committee for Scotland, CHI Advisory Group and the Privacy Advisory Committee of NHS National Services Scotland. The ethical and other permissions and related issues have been reported in detail [18,19] including an independent assessment by an ethicist.[21] Practice data custodians gave authorisation for the extraction, transfer and analysis of data for the study. All analysis outputs were subject to review by an NRS disclosure committee before release to researchers. NRS amended their statistical disclosure protocol during this study to round all cell sizes in statistical output to nearest multiple of five.

Results

Of the 17 practices approached, 10 agreed to take part in the study. Practice enrolment, information provision and obtaining approvals for data extraction was time-consuming and resource intensive.

The process of practice recruitment, data extraction and checking took the research team and Albasoft Ltd. approximately 12 months to complete. Five iterations of data extraction and check were required before all patients (including deregistered and deceased) and requested data fields were confirmed as being present in the database.

The primary care extract contained records for 110,698 people ever registered with the 10 practices up to 2011/2012 (Figure 1). 96,050 (86.8%) had an encrypted CHI number and of these 52,975 (55.2%) were successfully linked to the Census. Of these, around 8,500 (16.0%) belonged to non-White ethnic groups (Supplementary table 4) and 2,875 had an associated CVD event between May 2001 and April 2008. Investigation of unlinked primary care records showed 12,558 (11.3%) did not have an encrypted CHI number in our look-up table created in 2008 (i.e. registered with practices and were added to the CHI register after 2008). Of the 41,248 people with an encrypted CHI number, but no matching Census record, most (80.8%) were either born or had their first primary care record after the 2001 Census or were de-registered from the practice before 2001. Actual practice list sizes for 2001 were not available so for comparison purposes we calculated total list size (n=53,408) from 2011-2012 Quality & Outcomes Framework (QOF) statistics.[22]

Completeness of recording of cardiovascular risk factors and diabetes was broadly similar for most ethnic groups (Table 1). 89 people classified as having diabetes had a missing associated date in their primary care record. The crude prevalence of diabetes in White Scottish populations was 3.9% (women) and 5.2% (men). Indian, Pakistani and Other South Asian people had the highest crude rates. Figure 2 shows age-adjusted prevalence rate ratios for diabetes. Compared to White Scottish, prevalence was substantially higher for Pakistani populations: RR 274 [95%CI 238, 314] for men and 364 [305, 434] for women.

Within the linked cohort, 91.2% of records had a valid smoking status. Compared to White Scottish, most other ethnic groups had a similar level of recording of smoking status (Figure 2). For the CVD analysis, 82.5% (n=43,695) had smoking status recorded between 2001 and 2008. 64.7% were categorised as ‘never’ smokers and 35.3% as ‘ever’ smokers. Reported smoking was uncommon in Pakistani women (6.4%), and lower in Pakistani men (37.9%) compared to White Scottish men (50.3%).

There were 1700 (rounded to nearest 5, 53.8% in men) CVD events between 2001-2008 in the cohort of White Scottish, Other White British and Pakistani populations aged 30-74 with a valid smoking status (n=19,925 to nearest 5). Table 2 shows that age adjusted RRs of first hospitalisation or death for CVD were lower for Other White British men and women compared to White Scottish. Adjusting for education, smoking and diabetes did not substantially change the risks. Pakistani men and women had a higher risk of a first CVD event compared to White Scottish populations. For Pakistani men, adjusting for education and diabetes attenuated the risk. The RR for Pakistani women increased after

adjustment for education and smoking. Adjusting for diabetes attenuated the risk, but it still remained higher than in White Scottish women.

Discussion

Main findings of the study

This pilot study has demonstrated the processes and challenges involved in extracting a sample of risk factor and morbidity data from primary care and linking at individual level to national Census, hospital and mortality data. The recruitment of general practices and extraction of data was time consuming. However, our 58.8% uptake rate suggests practices engaged with the study objectives and the research team. The data extraction process required five iterations before a correct version was obtained. It proved technically challenging to understand the variations in coding systems for patient registration status among practices, whilst adhering to patient confidentiality protocols. In particular, information on deregistered and deceased patients was needed so that patients with retrospective data could be followed-up over time.

We judged the 55.2% linkage rate to be reasonable since practice lists included patients registered with the practice at any time point, while only those present in Scotland in 2001 could have participated in the Census and could be linked. Our final cohort of 52,975 people was comparable to total practice list size (n=53,408) from QOF statistics for 2011-2012.

Completeness levels in the primary care data extract were high for risk factors such as smoking status, weight and blood pressure. Incentivising recording through the QOF[23] is likely to have contributed to this. Furthermore, the risk factors we examined were of sufficient quality to interpret, categorise and utilise in risk model analysis.

Our pilot analysis using risk factors in models estimating ethnic variations in risk of CVD hospitalisation or death gave plausible results for the ethnic groups examined. Pakistani men and women had significantly higher age adjusted risks of a first CVD event compared to White Scottish, a finding consistent with previous reports in larger and more complete databases.[6,8-10] In men, further adjustment for education and diabetes attenuated this difference, although CVD risk remained elevated. Adjusting for education and smoking increased the already high risk of CVD for Pakistani-origin women, as expected, given the well described low prevalence of smoking in this group. The findings in relation to education were also broadly consistent with previous analyses[11] and those in relation to adjustment for diabetes and smoking had face validity. Our pilot study, therefore,

demonstrates both feasibility and potential value of this kind of linkage and supports the case for the creation and for the utility of national primary care datasets.

What is already known on this topic

In Scotland, as in many other countries, routine health service data collected in secondary care do not include information on clinical risk factors. However, this information is available from routinely recorded and coded electronic data from primary care systems. Within the context of general practice research databases in the UK and internationally[16] quality, completeness and use of risk factor data from primary care has been examined in detail, including for smoking[13,24] and diabetes.[25] Completeness of recording in the General Practice Research Database (GPRD) – now the Clinical Practice Research Datalink (CPRD) – was similar to our findings[24] and our results are also consistent with smoking prevalence estimates from the 2003 Scottish Health Survey.[26] The crude prevalence of 3.9-5.2% for diabetes is comparable with that reported by the 2005 Scottish Diabetes Survey (3.4%)[27] and the prevalence of diabetes was 3.7% in a large sample of Scottish practices.[25]

The methodologies for, and value of, linking primary care research databases to hospital and other disease register data are well documented. Williams et al summarised the issues in relation to data quality, new technology and methodology and models for linkage of disparate data sets.[12] With developments in informatics, computerisation and improved coding systems within primary care, de Lusignan et al suggested that linking primary care data on a large scale is likely to be a viable and efficient use of existing resources.[16]

Coding of patients' ethnic group is improving in both UK primary and secondary care records, but is still incomplete. Mathur et al reported a valid ethnic group code for 27% of patients in the 2012 build of the CPRD.[28] Ethnic coding in primary care records is believed to be similarly incomplete in Scotland (in 2010 an estimated 40% of patients had a valid ethnic code recorded in primary care within one large Scottish Health Board).[29] Therefore the best source of information on individual self-assigned ethnicity remains the Census. A New Zealand study examined consistency of ethnicity recording (amongst other variables) between primary and secondary care.[15] They reported high completeness and good quality recording of ethnic group in hospital data, but only 33% of general practice records recorded ethnicity and this was in a non-standardised format.

What this study adds

This study is, as far as we are aware, the first example of record linkage between primary care, Census, hospital and mortality records. It provides 'proof of concept' and valuable experience of such methods

and potential for further, larger scale research. It offers a method of studying inequalities including in ethnicity and health until routinely collected health databases contain good quality, complete, patient defined ethnic group, and other relevant characteristics, over a sufficient time period. We have shown that smoking and diabetes risk factor data recorded in our non-research general practices were of similar quality and completeness as reported from research practice databases.[24,25] A strength of our approach is the ability, potentially, to analyse by socio-demographic variables in the Census, such as sex and ethnic subgroups, and to incorporate major risk factor data from primary care records.

The current Scottish Primary Care Information Resource (SPIRE) initiative[30] in Scotland aims to create a national primary care dataset and the experience of this study illustrates the potential for the future exploitation of such data using record linkage.

Limitations of this study

This study was designed to test the feasibility of, and to pilot, accessing, linking and using risk factor information from primary care to strengthen analyses based on Census, mortality and secondary care data. We chose to explore published differences in risk of CVD among Scotland's ethnic minority populations.[6,8-10] Results in this paper are preliminary and should be treated with caution. Recruitment of only 10 selected general practices generated a sample which cannot be considered fully representative of the Scottish population.

Although we consider the 55.2% linkage rate of primary care to Census data as acceptable, we acknowledge the possibility of bias if non-linked Census records were not proportionately distributed amongst ethnic groups. With limited data available for non-linked records we were unable to study reasons for non-linkage for almost 19.2% (n=7902) of the primary care sample with an encrypted CHI number (Figure 1).

Time, resource and sample size constraints meant we examined only three ethnic groups for the CVD analysis and two clinical risk factors. Other clinical risk factor data from primary care, such as cholesterol measures or statin prescriptions, may not have similar levels of completeness and accuracy as smoking and diabetes. We cannot say definitively that our methods using CVD as an example would apply to other diseases and health conditions, but we think this is likely.

Conclusions

This study contributes new evidence on linkage of primary and secondary care and death data, to national Census records. Our methods offer the potential to examine disease risk at a national level, by sex, and ethnic group (given sufficient numbers), including separately different White populations and non-White ethnic groups. Our demonstration focussed on ethnicity but it is potentially valid for many

other Census variables. Linked national primary care datasets contain valuable information which may further our understanding of health inequalities and would be an important resource for public health and epidemiological surveillance and research. We hope our work will contribute to on-going efforts in primary care record linkage such as those proposed in the SPIRE project in Scotland, and beyond.

Funding

This work was supported by the Chief Scientist's Office (grant numbers CZH/4/ 648, CZH/4/878) and supplementary funding (no number) from NHS Health Scotland.

Acknowledgements

The authors are grateful for the support of the general practices who participated in the pilot study. Arti Nair, Jenny Holmes and Kath Ellis gave secretarial and administrative assistance during the study. Judith Fernandez contributed to the literature review and to preparation of the figures. Information Services Division (ISD) and National Records of Scotland (NRS) made many contributions to the work.

Authors' Contributions

The authorship, the authorship by-line, and note of contributions follows SHELS policy on authorship.

All authors served on the primary care subgroup of SHELS which planned the work in detail. Douglas coordinated the project during the second half of the study and was lead writer of this paper, Bhopal was the PI of SHELS, Fischbacher was a co-investigator and chair of the subgroup for part of the study, Simpson was a research fellow for the study and co-investigator, Steiner was clinical research fellow and analyst, Cézard was researcher and lead analyst, Bansal was the study coordinator and research fellow during the first half of the project, Sheikh is current chair of the subgroup, Ward is a collaborator and member of the subgroup. All authors helped plan the study and/or analysis, interpret data, critically revise drafts of the manuscript and agreed submission of the final draft.

Contributors from the Scottish Health and Ethnicity Linkage Study research team

These contributors served on the Steering Group and some on other important subgroups of SHELS, so gave general direction that helped this analysis. Chris Povey was a co-applicant and the originator

of the idea of linking the census data to the data held by ISD and he performed most of the linkage work to create the key look-up table. Prof Jamie Pearce (co-applicant) advised especially on socio-economic adjustment. Duncan Buchanan (co-applicant) chaired the analysis subgroup. Alex Stannard (part study) and Kirsty MacLachlan (part study) advised particularly in relation to NRS contributions. These important contributions did not meet ICMJE authorship requirements.

Data sharing

The data are not open access. Researchers interested in accessing these datasets should write to the Principal Investigator, Prof Raj Bhopal (raj.bhopal@ed.ac.uk).

Conflict of interests: None declared.

References

1 Scottish Government. Report on Health Inequalities 2015; http://www.scottish.parliament.uk/S4_HealthandSportCommittee/Reports/her-15-01w-rev.pdf. **accessed 15 Jul 2015**.

2 Anand SS, Yusuf S, Vuksan V *et al*. Differences in risk factors, atherosclerosis, and cardiovascular disease between ethnic groups in Canada: the Study of Health Assessment and Risk in Ethnic groups (SHARE). *Lancet* 2000; **356**(**9226**):279-284.

3 Yusuf S, Reddy S, Ounpuu S, Anand S. Global burden of cardiovascular diseases: part I: general considerations, the epidemiologic transition, risk factors, and impact of urbanization. *Circulation* 2001; **104**(**22**):2746-2753.

4 de Bruijne MC, van Rosse F, Uiters E *et al*. Ethnic variations in unplanned readmissions and excess length of hospital stay: a nationwide record-linked cohort study. *Eur J Public Health* 2013; **23**(**6**):964-971.

5 Jatrana S, Blakely T. Ethnic inequalities in mortality among the elderly in New Zealand. *Aust N Z J Public Health* 2008; **32**(**5**):437-443.

6 Bansal N, Fischbacher CM, Bhopal RS *et al*. Myocardial infarction incidence and survival by ethnic group: Scottish Health and Ethnicity Linkage retrospective cohort study. *BMJ Open* 2013; **3**(**9**):e003415.

7 Deedwania P. Diabetes, cardiovascular disease, and ethnicity. *J Am Coll Cardiol* 2013; **61**(**17**):1787-1789.

8 Bhopal RS, Bansal N, Fischbacher CM, Brown H, Capewell S. Ethnic variations in heart failure: Scottish Health and Ethnicity Linkage Study (SHELS). *Heart* 2012; **98**(**6**):468-473.

9 Bhopal RS, Bansal N, Fischbacher C, Brown H, Capewell S. Ethnic variations in chest pain and angina in men and women: Scottish Ethnicity and Health Linkage Study of 4.65 million people. *Eur J Prev Cardiol* 2012; **19**(**6**):1250-1257.

10 Bhopal RS, Bansal N, Fischbacher CM, Brown H, Capewell S. Ethnic variations in the incidence and mortality of stroke in the Scottish Health and Ethnicity Linkage Study of 4.65 million people. *Eur J Prev Cardiol* 2012; **19**(**6**):1503-1508.

11 Fischbacher CM, Cezard G, Bhopal RS, Pearce J, Bansal N. Measures of socioeconomic position are not consistently associated with ethnic differences in cardiovascular disease in Scotland: methods from the Scottish Health and Ethnicity Linkage Study (SHELS). *Int J Epidemiol* 2014; **43**(**1**):129-139.

12 Williams T, van ST, Puri S, Eaton S. Recent advances in the utility and use of the General Practice Research Database as an example of a UK Primary Care Data resource. *Ther Adv Drug Saf* 2012; **3**(**2**):89-99.

13 Simpson CR, Hippisley-Cox J, Sheikh A. Trends in the epidemiology of smoking recorded in UK general practice. *Br J Gen Pract* 2010; **60**(**572**):e121-e127.

14 Martin-Merino E, Ruigomez A, Johansson S, Garcia-Rodriguez LA. Hospitalised ischaemic cerebrovascular accident and risk factors in a primary care database. *Pharmacoepidemiol Drug Saf* 2011; **20**(**10**):1050-1056.

15 Tomlin A, Hall J. Linking primary and secondary healthcare databases in New Zealand. *N Z Med J* 2004; **117**(**1191**):U816.

16 de Lusignan S, van Weel C. The use of routinely collected computer data for research in primary care: opportunities and challenges. *Fam Pract* 2006; **23**(**2**):253-263.

17 McCartney M. Care.data: why are Scotland and Wales doing it differently? *BMJ* 2014; **348**:g1702.

18 Bhopal R, Fischbacher C, Povey C *et al*. Cohort profile: Scottish health and ethnicity linkage study of 4.65 million people exploring ethnic variations in disease in Scotland. *Int J Epidemiol* 2011; **40**(**5**):1168-1175.

19 Fischbacher CM, Bhopal R, Povey C *et al*. Record linked retrospective cohort study of 4.6 million people exploring ethnic variations in disease: myocardial infarction in South Asians. *BMC Public Health* 2007; **7**:142.

20 Chisholm J. The Read clinical classification. *BMJ* 1990; **300**(**6732**):1092.

21 Boyd KM. Ethnicity and the ethics of data linkage. *BMC Public Health* 2007; **7**:318.

22 Information Services Division. Quality and Outcomes Framework 2011-12. https://isdscotland.scot.nhs.uk/Health-Topics/General-Practice/Quality-And-Outcomes-Framework/2011-12/practicelevel-summaries.asp. **accessed 27th July 2015**.

23 British Medical Association. QOF guidance 2015-2016 according to nation http://bma.org.uk/practical-support-at-work/contracts/gp-contracts-and-funding/independent-contractors/qof-guidance. **accessed 15 July 2015**.

24 Delaney JA, Daskalopoulou SS, Brophy JM *et al*. Lifestyle variables and the risk of myocardial infarction in the general practice research database. *BMC Cardiovasc Disord* 2007; **7**:38.

25 McGovern MP, Williams DJ, Hannaford PC *et al*. Introduction of a new incentive and target-based contract for family physicians in the UK: good for older patients with diabetes but less good for women? *Diabet Med* 2008; **25**(**9**):1083-1089.

26 Scottish Government. Scottish Health Survey 2003. http://www.gov.scot/Publications/2005/12/02160336/03493. **accessed 15 July 2015**.

27 Scottish Diabetes Group. Scottish Diabetes Survey 2010. http://www.diabetesinscotland.org.uk/Publications.aspx?catId=3. **accessed 15 July 2015**.

28 Mathur R, Bhaskaran K, Chaturvedi N *et al*. Completeness and usability of ethnicity data in UK-based primary care and hospital databases. *J Public Health (Oxf)* 2014; **36**(**4**):684-692.

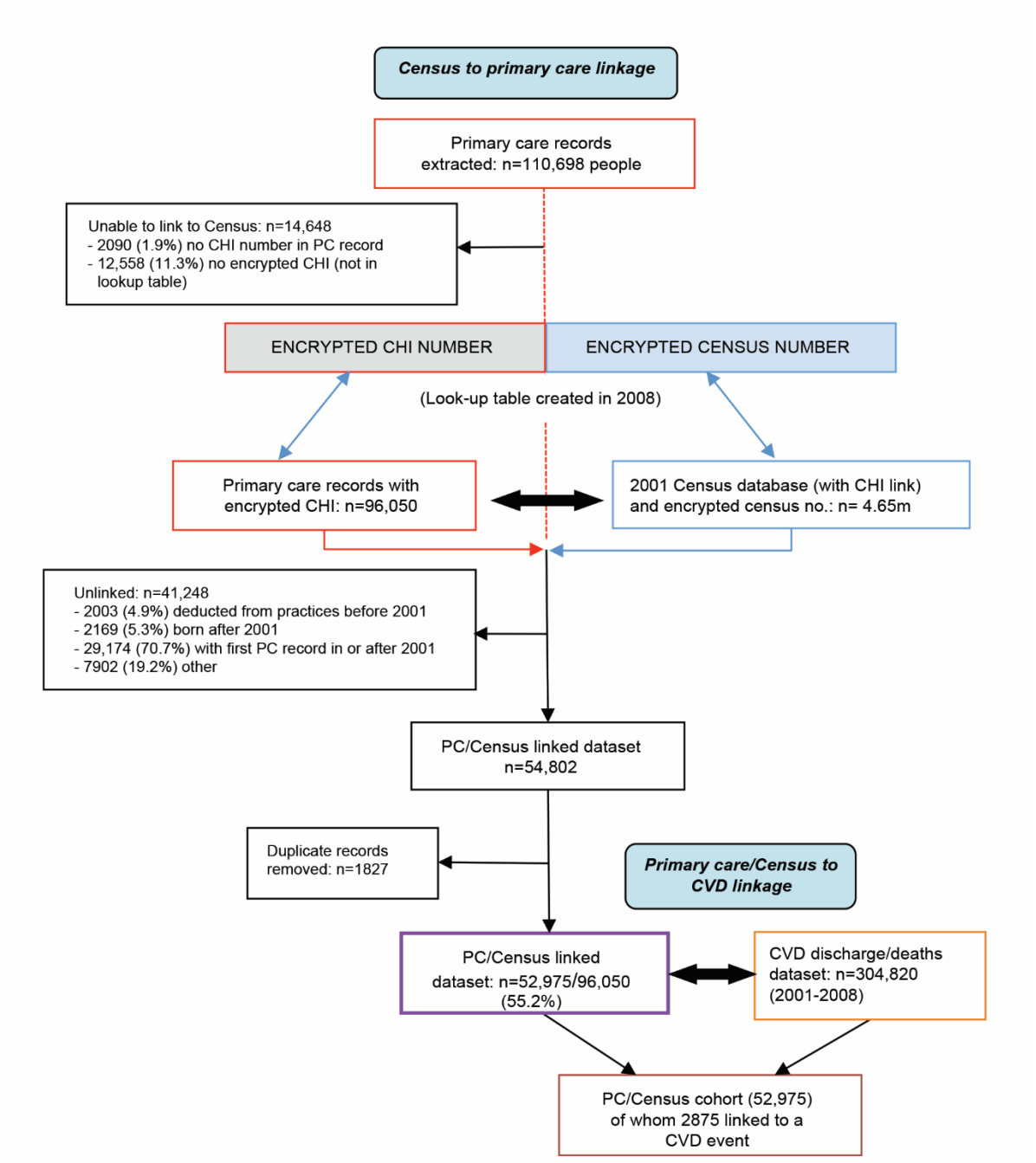
29 NHS Lothian and University of Edinburgh. NHS Lothian Ethnicity Monitoring case study May 2013. http://www.healthscotland.com/documents/21006.aspx. **accessed 15 July 2015**.

30 NHS Scotland. Scottish Primary Care Information Resource (SPIRE) 2015. http://www.spire.scot.nhs.uk/. **accessed 15 July 2015**.

Box 1 Scotland 2001 Census ethnic groups (self-defined)

White Scottish
Other White British
White Irish
Other White
Any mixed background
Indian
Pakistani
Bangladeshi
Other South Asian
Caribbean
African
Black Scottish or Other Black
Chinese
Other ethnic group

Figure 1 Linkage of primary care to Census data and to cardiovascular hospitalisation/deaths



Notes: PC - Primary Care

Table 1. Completeness/availability of primary care risk factor information (%) for people in the linked primary care to Census cohort

Ethnic group	N	Smoking information	Weight	BMI	Blood Pressure	Cholesterol	Diabetes[a]
MEN		%	%	%	%	%	%
White Scottish	18104	90.2	69.2	53.6	72.6	30.4	5.2
Other White British	1727	89.5	73.5	53.3	68.4	24.3	3.9
White Irish	397	87.9	69.3	50.4	79.1	34.3	4.8
Other White	621	86.3	72.8	50.1	66.2	25.4	2.9
Indian	627	89.2	75.8	53.8	79.4	45.6	15.0
Pakistani	3008	86.1	71.7	52.4	78.2	42.2	14.4
Other South Asian	258	85.3	72.5	51.9	70.5	35.7	10.5
Chinese	176	86.9	73.3	47.7	72.7	26.1	3.4
African origin	101	90.1	79.2	52.5	86.1	28.7	*
WOMEN							
White Scottish	21670	91.2	73.5	56.5	84.0	29.2	3.9
Other White British	2114	90.8	76.2	57.2	79.8	20.6	2.8
White Irish	500	88.8	73.0	49.0	83.4	31.0	2.4
Other White	815	91.4	76.4	52.3	76.2	19.4	1.8
Indian	601	88.0	77.4	51.4	86.5	42.4	8.7
Pakistani	3006	87.9	75.1	53.1	83.2	40.6	12.3
Other South Asian	230	87.4	77.0	54.4	81.7	30.4	9.6
Chinese	198	84.9	76.3	50.0	71.8	21.8	*
African origin	78	83.3	71.8	54.0	84.3	23.7	*

[a] A record of diabetes equates to crude diabetes prevalence rates *Data not released from NRS because of the risk of disclosure of personal information.

Figure 2 Prevalence ratios for a record of diabetes and presence of smoking information in primary care dataset, by ethnic group and sex

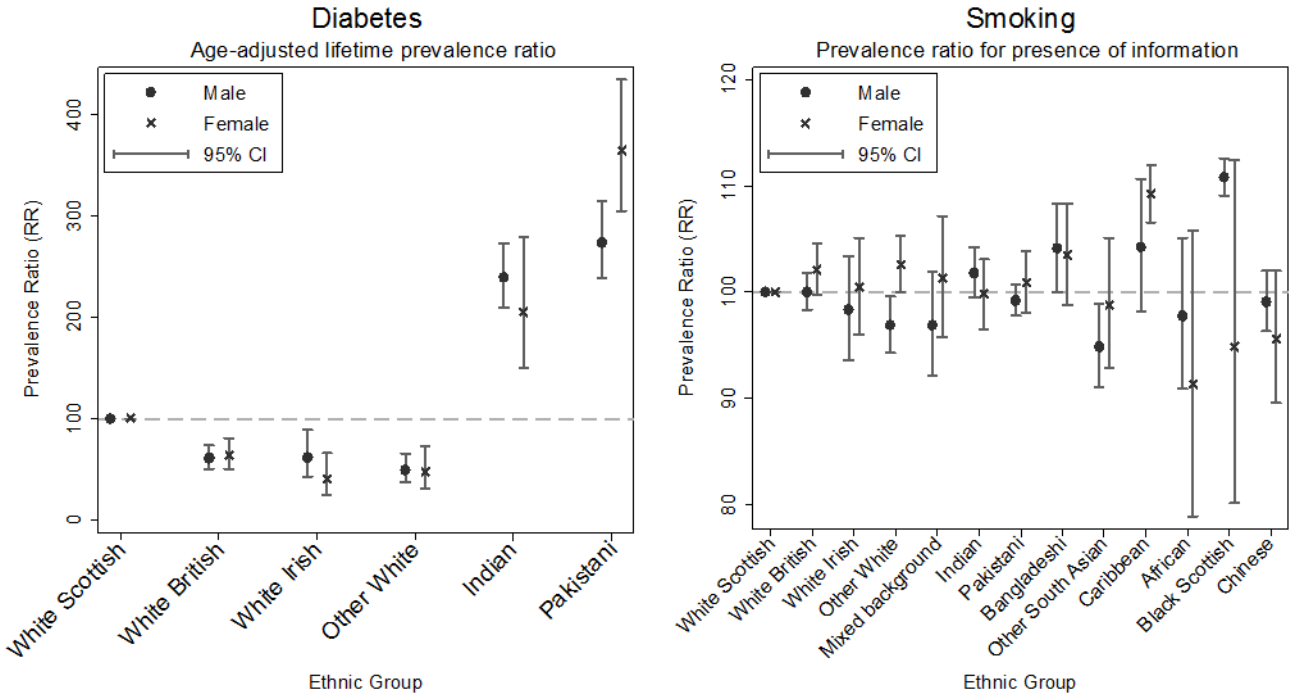


Table 2. Age adjusted Poisson rates per 100,000 person years (PY) and relative risks (RR) for first CVD hospitalisation/death between 2001-2008 by sex, for three ethnic groups. RRs are adjusted by age, highest education level, smoking status and diabetes with 95% confidence intervals (CI)

Ethnic group	Any first CVD event[a]	PY[a]	Rate (for 100,000 PY)	Poisson relative risks (RR and 95%CI) adjusted by:				
MALE				**Age**	**Age and education**	**Age, education and smoking**	**Age, education and diabetes**	**Age, education smoking & diabetes**
White Scottish	740	49950	1477.5	100.0	100.0	100.0	100.0	100.0
Other White British	45	4730	1089.4	73.7 (48.6, 111.8)	84.6 (56.5, 126.5)	86.3 (58.5, 127.3)	83.9 (56.1, 125.5)	85.6 (58.0, 126.4)
Pakistani	130	6760	2262.8	153.1 (121.7, 192.6)	139.3 (111.3, 174.5)	150.9 (122.7, 185.7)	121.1 (96.4, 152.0)	132.5 (106.6,164.6)
FEMALE								
White Scottish	650	58670	1107.9	100.0	100.0	100.0	100.0	100.0
Other White British	30	4845	731.6	66.0 (42.0, 103.7)	77.3 (49.4, 121.1)	80.5 (52.0, 124.5)	77.6 (49.2 122.3)	80.7 (51.9, 125.5)
Pakistani	105	6580	2192.5	197.9 (148.8, 263.2)	159.3 (127.3, 199.5)	202.2 (166.4, 245.7)	137.5 (112.4, 168.3)	174.3 (143.8, 211.1)

[a] rounded to nearest 5