# Use of laboratory-based surveillance data to estimate the number of people chronically infected with hepatitis B living in Scotland

Schnier, Christian; Wallace, Lesley; Templeton, Kate; Aitken, Celia; Gunson, Rory; Molyneaux, Pamela; McIntyre, Paul; Povey, Christopher ; Goldberg, David; Hutchinson, Sharon

**Use of laboratory-based surveillance data to estimate the number of people chronically infected with hepatitis B living in Scotland**

C. Schnier[1,2], L. Wallace[2], K. Tempelton[3], C. Aitken[4], R.N. Gunson[4], P. Molyneaux[5], P. McIntyre[6], C. Povey[7], D. Goldberg[2], S. Hutchinson[1,2]

[1]School of Health and Life Sciences, Glasgow Caledonian University, Glasgow G4 0BA, UK

[2]Health Protection Scotland, Meridian Court, 5 Cadogan Street, Glasgow G2 6QE, UK

[3]Specialist Virology Centre, Royal Infirmary of Edinburgh, Edinburgh EH16 4SA, UK

[4]West of Scotland Specialist Virology Centre, Gartnaval General Hospital, Glasgow G12 0YN, UK

[5]Department of Medical Microbiology, Aberdeen Royal Infirmary, Aberdeen AB25 2ZN, UK

[6]Department of Medical Microbiology, Ninewells Hospital & Medical School, Dundee DD1 9SY, UK

[7]Information Service Division, Scotland, 1 South Gyle Crescent, Edinburgh EH12 9EB, UK


**Corresponding author & Requests for Reprints**

C. Schnier

School of Health and Life Sciences

Glasgow Caledonian University

Glasgow, G4 0BA

Email: christian.schnier@nhs.net


   Tel: +44 141 282 2956

Fax: +44 141 300 1170

**Short running head**   Hepatitis B prevalence in Scotland

# ABSTRACT

It is paramount to understand the epidemiology of chronic Hepatitis B to inform national policies on vaccination and screening/testing as well as cost-effectiveness studies. However, information on the national (Scottish) prevalence of chronic Hepatitis B by ethnic group is lacking. To estimate the number of people with chronic Hepatitis B in Scotland in 2009 by ethnicity, gender and age, the test data from virology laboratories in the four largest cities in Scotland **were** combined with estimates of the ethnic distribution of the Scottish population. **Ethnicity in both the test data and the Scottish population was derived using a name-based ethnicity classification software (Onomap)**. **For 2009, we estimated** 8720 (95% confidence interval: 7490 to 10230) people aged 15 and older were living with chronic hepatitis B infection in Scotland. This corresponds to 0.2% (CI: 0.17% to 0.24%) of the Scottish population aged 15 and older. **Although** East and South Asians make up a small proportion of the Scottish population, they make up 44% of the infected population. In addition, 75% of those infected are aged 15-44 and almost 60% are male. **This study quantifies for the first time on a national level the burden of chronic hepatitis B infection by ethnicity, gender and age.** It confirms the importance of promoting and targeting ethnic minority groups for hepatitis B testing.

# INTRODUCTION

Worldwide, it has been estimated that about 2 billion people have been infected with hepatitis B virus (HBV). 240 million have chronic HBV infection (CHB) and 600000 die each year either from CHB-associated liver cirrhosis or hepato-cellular carcinoma [1]. A UK prevalence of 0.3% for CHB has been reported by the Department of Health [2]. This estimate relates to data from antenatal screening tests in the West Midlands between 1983 and 1985 [3]. More recently, a higher prevalence of 0.45% has been predicted for England and Wales as a result of modeling data on notifications of acute HBV [4]. Generally, higher prevalences of CHB have been found amongst men and older people [5] and in members of certain ethnic minority groups, especially South Asians and East Asians [6], [7].

In the UK, universal infant or adolescent immunisation is not considered cost-effective [8]. Instead, the joint Committee on Vaccination and Immunization has recommended that HBV immunisation should be targeted at individuals at high risk of HBV exposure and at people at increased risk of complications should they acquire HBV [9].

Early diagnosis and treatment of HBV is cost-effective [10], but identification of individuals, who are often asymptomatic, is difficult. In The Netherlands and Canada, both countries with a low HBV prevalence [11], screening first generation migrants and early treatment of those with CHB has been shown to be cost-effective [12], [13]. HBV testing has been part of the UK antenatal screening programme since 1998 [14]. Recent guidance from the National Institute for Health and Clinical Excellence (NICE) recommends testing people at increased risk of HBV infection. Those at higher risk are particularly migrants from medium- or high-prevalence countries, people who inject or have injected drugs and men who have sex with men [15].

Until now, most studies on HBV prevalence have focussed on the tested population without comparisons made to the general population. In this study,

we have developed an approach based on a combination of HBV laboratory test data with estimates of the Scottish population as a means to estimate, by major ethnic groups, the number of people living with CHB in Scotland. This approach is of paramount importance to guide national policy and screening strategies. **Nevertheless, despite it not being a resource intensive approach, it has not hitherto been implemented.**

# METHODS

Laboratory HBV test data were used to determine the prevalence of CHB by ethnic background in (i) women undergoing an antenatal screen, and (ii) women and men tested in other primary and secondary care settings. The former gives information on the prevalence of CHB in females of childbearing age, while the latter gives information on differences in the prevalence of CHB by age and gender. To generate an estimate of the prevalence of CHB infection among adults living in Scotland, estimates of CHB prevalence by age, gender and ethnicity were then combined with estimates of the general Scottish population.

## Laboratory data

The virology laboratories in the four largest cities in Scotland (Aberdeen, Dundee, Edinburgh and Glasgow) provided Health Protection Scotland (HPS) with data on all HBV-related diagnostic tests conducted at their laboratories from $1^{st}$ January 2005 to $31^{st}$ December 2009 on samples submitted from their respective health board areas. Those areas represent approximately 60% of the Scottish population. Information for each sample included diagnostic test results by date, patient identifiers (including full names, date of birth, gender, postcode of residence and the community health index (CHI) number when available [16]), referral source of test (i.e. primary care, secondary care, antenatal and other), and referral source/clinic number; **information on ethnicity is not routinely reported with a virology test request and therefore could not**

**be relied on for analysis.**

Hepatitis B surface antigen (HBsAg)-positive tests and HBsAg-negative tests were used to define infected and uninfected individuals, respectively. Indeterminate, weak positive or inconclusive results were excluded. In order to ensure that HBV infection included in the study was CHB, an assessment of HBV positive individuals was performed. An HBsAg positive individual was excluded if there was (1) a negative test within 180 days before the first positive HBsAg-test or (2) a positive test for IgM antibodies to the hepatitis B core antigen at the time of the first positive HBsAg-test.

Ethnicity was assigned using a name-based 'ethnicity' classification methodology (Onomap). The Onomap software classifies ethnicity based on a comparison of forename and surname against a database of names sourced from public name registries from over 26 countries. In order to correct for misclassification in the assignment of ethnicity, predictive values from a recent validation study of Onomap have been used [16]. This study, set in the Scottish population and involving large datasets of birth registrations and the pupil census, concluded that Onomap was an effective methodology for categorizing populations into a variety of ethnic groups. When comparing predictions from Onomap to birth registrations, the study found predictive values to determine British ethnicity of 94.6% (positive predictive value, PPV; **the proportion of people determined by Onomap to belong to a certain ethnic group which truly belong to that ethnic group**) and 74.4% (negative predictive value, NPV; **the proportion of people determined by Onomap to not belong to a certain ethnic group which truly don't belong to that ethnic group**), Chinese ethnicity of 70.9% PPV and 99.9% NPV and South Asian ethnicity of 53.6% PPV and 99.5% NPV. South Asian countries include Iran, Afghanistan, Pakistan, Tibet, Nepal, India, Bangladesh, Burma and Sri Lanka. East Asian countries include Mongolia, the People's Republic of China, Japan, North Korea, South Korea and Taiwan. In accordance with data protection requirements, after classification of ethnicity surnames were replaced by surname Soundex (a

consonant-only phonetic encoding) and forenames were replaced with forename initials. **HPS Clinical Governance approval has been given for use of names to assign ethnicity.**

Records for the same person were identified using (i) a combination of forename initial, soundex of surname, gender and date of birth, (ii) the CHI number, or (iii) the referral source clinic-number. Age was calculated at the time of the first HBsAg test. To avoid problems from a low number of tests in children under the age of 15, analysis was restricted to the adult population (here regarded as those aged 15 years or older).

## HBV prevalence in women aged 15-44 years, by ethnicity

The prevalence of HBV infection in women aged 15-44 years was estimated using test results from women **living in either of the four NHS board areas** and who had been tested for HBsAg as part of the universal antenatal screening programme. It was assumed that pregnant women tested as part of this programme were broadly representative of all women of child-bearing age, since very few women, about 2% in Glasgow in 2004-2009 [17], opt out of antenatal screening.

154927 HBsAg tests on the laboratory database were recognized as antenatal tests. Although it was not possible to identify all antenatal tests in the database, it was assumed that misclassification in the database was non-differential.

Prevalence was stratified according to age (15-29 and 30-44 years) and ethnic group (British and other, South Asian, East Asian and unknown). To avoid bias from multiple testing, only the first HBsAg test for every woman during the period 2005-2009 was used in the analysis, leaving in total 129171 tests for analysis.

## HBV prevalence in women older than 44 years and men older than 14 years, by ethnicity

CHB prevalence in women older than 44 years and men older than 14 years was estimated using as a starting point the estimated prevalence generated for women aged 15-44 years (as described above). Estimates were then adjusted according to relative differences in prevalence for the former groups, as determined from examination of data on all men and women tested for HBsAg in primary and secondary health care settings (excluding antenatal screens).

96763 HBsAg tests on the laboratory database had referral source identified as either primary (29393) or secondary (67370) health care provider (excluding tests from Genito-urinary medicine providers, occupational health services and renal screening) and where the test was not also recognized as an antenatal screening test. Again, only the first HBsAg test for individuals was used in the analysis. For each ethnic group separately, Poisson regression was used to generate relative risks of testing HBsAg-positive for (a) men compared to women, and (b) women older than 44 years compared to those aged 15-29 years, adjusting for NHS board and referral source. For East Asians, women aged 15-44 were grouped together to avoid problems with small numbers in the Poisson models. Year of test (i.e. 2005, 2006, 2007, 2008, 2009) was not a significant confounder of HBsAg positivity and therefore was not included further in the regression models. Similarly, no significant interaction between age and gender was found and the interaction term was therefore not included in regression models.

For each ethnic group, the estimated prevalence of HBsAg positivity in women older than 44 years was calculated by multiplying the prevalence in women aged 15-29 years (generated from the antenatal test data) with the relative risk of testing HBsAg-positive among women older than 44 years compared to those aged 15-29 years (generated from primary and secondary health care test data). The prevalence in men was similarly estimated, applying the relative

risk for men compared to women. For example, to estimate the prevalence in British men older than 44 years, the prevalence in British women, aged 15-29 years (0.16%, Table 1) was multiplied with the adjusted relative risk of British men and women older than 44 year compared to British men and women aged 15-29 years (RR 0.3, Table 2). The resulting product (0.05%) was then multiplied with the adjusted relative risk of British men compared to British women (RR 1.69, Table 2) resulting in an estimated prevalence of 0.08% (Table 4). In estimating the ethnic-specific CHB prevalences in older women and men, it was assumed that the relative risks, as calculated from tests conducted in primary and secondary health care settings, were representative of differences in the general population.

## Scotland's population in 2009, by ethnicity

Names, gender and age of all people registered with a general practitioner (GP) in Scotland in June 2011 (N=5800674 observations) were extracted from the CHI database at Information Service Division, Scotland. **The number of people in the CHI database slightly exceeds the number of people living in Scotland because it can include some visitors, duplicate records, people who have moved out of Scotland (and are not known to have moved) and people who have died in that year.** The distribution of the Scottish populations by age (15-29, 30-44 and > 44), gender and ethnicity (British and Other, South Asian and East Asian) in 2011 was then estimated using Onomap. To correct for misclassification in the assignment of ethnicity, the estimated distribution was then corrected using published predicted values [16]. Finally, corrected estimates were applied to the size of the Scottish population in 2009 according to age group and gender, predicted by the General Register Office for Scotland (GRO; **N=5194000** [18]). **In using data from CHI, it was assumed that ethnic minority groups are represented in the CHI database in the same proportion as in the general population.** Estimates generated from the CHI database were compared to results

from the latest (2001) census **(N=5062000)** [19]**, at the time of analysis,** and to published estimates **for 2010** by Wohland *et al.* **(N=4940000)** [20], who predicted the size of the Scottish population using estimates of ethnic group fertility, ethnic group mortality, internal and international migration.

## Scotland's CHB population in 2009, by ethnicity

Scotland's CHB population in 2009 was estimated as the product of the estimated CHB prevalence and the estimated size of the population in 2009 by **gender,** age group and ethnicity. Confidence intervals (CI) for these estimates were derived by replacing the following three parameters with distributions and then repeatedly (10000 iterations) randomly sampling from those distributions: i) the prevalence of CHB in the antenatal population, ii) the log of the adjusted relative risk of testing HBsAg-positive among women older than 44 years compared to those aged 15-29 years and men compared to women and iii) the predictive values for Onomap classification. For each ethnic group separately, parameters were replaced with the following distributions: i) the prevalence of CHB in the antenatal population was replaced with a binomial distribution with n equal to the number of women tested in antenatal tests and p equal to the proportion of n who tested positive; ii) the log of the adjusted relative risks was replaced with a normal distribution with mean of the log of the adjusted RR and the standard error (SE) as estimated from the Poisson models; iii) the predictive values for Onomap classification was replaced with a normal distribution with mean and the SE as published by [16]. In generalizing results to the Scottish population, it was assumed that the ethnicity, gender and age-specific CHB prevalence estimates for the four largest Scottish NHS boards was representative of the Scottish population.

## Data processing and analysis

PostgreSQL (Version 9.0.1, PostgreSQL Global Development Group) was used for data storage and data linkage. R statistical package (V 2.13.1, R Foundation

Table 1: Number of women tested for HBsAg and percentage testing HBsAg-positive at their first recorded antenatal screening between 2005 and 2009 in Scotland.

| Age group | N | % | HBsAg+ve | .% of N (95% CI) |
|---|---|---|---|---|
| British and other | | | | . |
| 15-29 | 64397 | 50 | 103 | 0.16 (0.13 to 0.19) |
| 30-44 | 55006 | 43 | 55 | 0.10 (0.08 to 0.13) |
| Both age groups | 119403 | 92 | 158 | 0.13 (0.11 to 0.15) |
| South Asian | | | | . |
| 15-29 | 3518 | 3 | 15 | 0.43 (0.24 to 0.70) |
| 30-44 | 2117 | 2 | 18 | 0.85 (0.50 to 1.34) |
| Both age groups | 5635 | 4 | 33 | 0.59 (0.40 to 0.82) |
| East Asian | | | | . |
| **15-29** | **841** | **0.6** | **106** | **12.60 (10.44 to 15.04)** |
| **30-44** | **687** | **0.4** | **51** | **7.42 (5.58 to 9.65)** |
| Both age groups | 1528 | 1 | 157 | 10.27 (8.80 to 11.91) |
| Unknown | | | | . |
| 15-29 | 1680 | 1 | 10 | 0.60 (0.29 to 1.09) |
| 30-44 | 925 | 1 | 17 | 1.84 (1.07 to 2.93) |
| Both age groups | 2605 | 2 | 27 | 1.04 (0.68 to 1.50) |
| All ethnic groups | | | | . |
| 15-44 | 129171 | 100 | 375 | 0.29 (0.26 to 0.32) |

for Statistical Computing) and @RISK (V5, Palisade Corporation) were used for statistical analysis and modelling.

# RESULTS

## HBV prevalence in women aged 15-44 years, by ethnicity

375 women tested HBsAg-positive in 129171 antenatal screening tests in Scotland during 2005-2009 (Table 1). This corresponds to a prevalence of 0.29% (95% CI 0.26-0.32%). The majority of women for whom an ethnic group could be determined were of British and Other ethnicity (94%). The prevalence in this group was 0.13%. Higher proportions of women in the South Asian (0.59%) and East Asian (10.3%) ethnic groups were found to be HBsAg positive. Among South Asian women, the prevalence was higher for women aged 30-44 compared to 15-29 years, whereas among women with **East Asian and** British and Other ethnicity the prevalence of HBsAg was higher for women aged 15-29.

## HBV prevalence in women older than 44 years and men older than 14 years, by ethnicity

Compared to females of South Asian and British and Other ethnicity, males had a 1.7 - 1.8 times higher risk of testing HBsAg-positive (Table 2). No significant gender difference in risk was observed for East Asian ethnicity. Compared to those aged 15-29, those aged 44 years and older of British and Other ethnicity had a significantly lower risk of testing HBsAg-positive (RR 0.3, 95%CI 0.2 to 0.4). No significant difference in risk by age was observed for those of South and East Asian ethnicity. **Compared to women and men tested in secondary care, those tested in primary care had a higher risk of testing positive. This pattern was observed for all ethnic groups and age groups with the exception of 30-44 year old East Asian females who had a similar risk in primary and secondary care.**

Multiplying the adjusted relative risk estimates for each ethnic group (Table 2) with the prevalences of HBsAg positivity in women aged 15-29 years (Table 1), estimates of the prevalence of CHB were generated (Table 4(a)). Estimates for the prevalence of CHB in men ranged from 0.08% in those of British/Other ethnicity aged over 44 years to 11.6% in those of East Asian ethnicity aged 15-44 years. Estimates for the prevalence of CHB in women aged over 44 years ranged from 0.05% in those of British and Other ethnicity to 9.3% in those of East Asian ethnicity.

## Scotland's population in 2009, by ethnicity

Through the approach of applying Onomap to the CHI database it was estimated that 1.5% and 0.7% of the Scottish population aged 15 years or older in 2009 were of South Asian and East Asian ethnicity, respectively (Table 3). For South Asians, the estimate was higher than that reported from the 2001 Sottish Census (1.0%) and from [20] for the year 2010 (1.1%). For East Asians, it was higher than that reported from the 2001 census (0.3%), but comparable to the

Table 2: Relative risks of testing HBsAg-positive among men and women older than 15 years who had been tested for HBsAg in primary and secondary care settings (excluding all antenatal tests) in Scotland during 2005-2009, stratified by ethnicity.

| Variable | Level | N | % | HBsAg-pos | % of N | RR unadjusted | adjusted[a] |
|---|---|---|---|---|---|---|---|
| (a) British and other | | | | | | | |
| Source | Secondary Care | 53957 | 68 | 271 | 0.50 | | |
| | Primary Care | 25768 | 32 | 226 | 0.88 | 1.75 (1.46 to 2.08) | 1.57 (1.31 to 1.88) |
| Gender | Female | 36546 | 46 | 166 | 0.45 | | |
| | Male | 43179 | 54 | 331 | 0.77 | 1.69 (1.4 to 2.03) | 1.69 (1.40 to 2.04) |
| Age group | 15-29 | 15533 | 19 | 160 | 1.03 | | |
| | 30-44 | 24359 | 31 | 219 | 0.90 | 0.87 (0.71 to 1.07) | 0.83 (0.68 to 1.02) |
| | >44 | 39833 | 50 | 118 | 0.30 | 0.29 (0.23 to 0.36) | 0.30 (0.24 to 0.39) |
| NHSboard | GG&C[b] | 30762 | 39 | 248 | 0.81 | | |
| | Other | 48963 | 61 | 249 | 0.51 | 0.63 (0.53 to 0.75) | 0.66 (0.55 to 0.79) |
| (b) South Asian | | | | | | | |
| Referral | Secondary Care | 1431 | 44 | 30 | 2.10 | | |
| Source | Primary Care | 1830 | 56 | 83 | 4.54 | 2.16 (1.43 to 3.28) | 2.05 (1.34 to 3.13) |
| Gender | Female | 1159 | 36 | 25 | 2.16 | | |
| | Male | 2102 | 64 | 88 | 4.19 | 1.94 (1.24 to 3.03) | 1.81 (1.16 to 2.83) |
| Age group | 15-29 | 1127 | 35 | 33 | 2.93 | | |
| | 30-44 | 1287 | 39 | 50 | 3.89 | 1.33 (0.85 to 2.06) | 1.25 (0.81 to 1.95) |
| | >44 | 847 | 26 | 30 | 3.54 | 1.21 (0.74 to 1.98) | 1.25 (0.76 to 2.06) |
| NHSboard | GG&C | 2056 | 63 | 75 | 3.65 | | |
| | Other | 1205 | 37 | 38 | 3.15 | 0.86 (0.59 to 1.28) | 0.95 (0.64 to 1.40) |
| (c) East Asian | | | | | | | |
| Referral | Secondary Care | 503 | 45 | 101 | 20.08 | | |
| Source | Primary Care | 627 | 55 | 146 | 23.29 | 1.16 (0.90 to 1.49) | 1.22 (0.94 to 1.57) |
| Gender | Female | 551 | 49 | 110 | 19.96 | | |
| | Male | 579 | 51 | 137 | 23.66 | 1.19 (0.92 to 1.52) | 1.13 (0.88 to 1.46) |
| Age group | 15-44 | 887 | 78 | 198 | 22.32 | | |
| | >44 | 243 | 22 | 49 | 20.16 | 0.90 (0.66 to 1.23) | 0.91 (0.66 to 1.24) |
| NHSboard | GG&C | 489 | 43 | 134 | 27.40 | | |
| | Other | 641 | 57 | 113 | 17.63 | 0.64 (0.50 to 0.83) | 0.63 (0.49 to 0.82) |

[a]Adjusted for Source, Gender, Age and NHS board

[b]Greater Glasgow & Clyde

Table 3: Distribution (percentage) of the Scottish population by ethnicity, sex and age group. a) CHI database (adjusted to 2009 population), b) ETHPOP estimate for the population in 2010 and c) Census in 2001

| Age group | a) CHI database[a] (2009) | | b) ETHPOP estimate (2010) | | c) Census (2001) | |
|---|---|---|---|---|---|---|
| | Females | Males | Females | Males | Females | Males |
| British and other | | | | | | |
| 15-29 | 11.17 | 11.54 | 10.94 | 11.41 | 11.22 | 11.10 |
| 30-44 | 12.10 | 11.26 | 11.89 | 11.18 | 14.22 | 13.34 |
| >44 | 27.87 | 23.85 | 28.52 | 24.29 | 26.59 | 22.26 |
| South Asian | | | | | | |
| 15-29 | 0.24 | 0.36 | 0.18 | 0.19 | 0.19 | 0.20 |
| 30-44 | 0.23 | 0.34 | 0.18 | 0.19 | 0.15 | 0.17 |
| >44 | 0.15 | 0.20 | 0.17 | 0.19 | 0.11 | 0.13 |
| East Asian | | | | | | |
| 15-44 | 0.28 | 0.26 | 0.25 | 0.25 | 0.12 | 0.12 |
| >44 | 0.07 | 0.08 | 0.10 | 0.09 | 0.04 | 0.04 |

[a]Adjusted for PPV and NPV of name classification using Onomap and for the predicted size of the population.

estimate from [20](0.7%).

## Scotland's CHB population in 2009, by ethnicity

By applying the CHB prevalence estimates (Table 4a) to the estimates of the adult Scottish population by ethnicity, gender and age group (as in Table 3a) it was estimated that approximately 8720 (95% CI 7490-10230) people aged 15 and older infected with CHB were living in Scotland in 2009; this represents a population prevalence among people aged 15 and older of 0.2% (95% CI 0.17-0.24%). Of the 8720, 56% (4920, CI 3970-6050) were predicted to have British and Other ethnicity, 37% (3200, CI 2460-4120) to have East Asian ethnicity and 7% (600 CI 360-960) to have South Asian ethnicity. 75% (6530 CI 5630-7580) were predicted to be between 15 and 44 years old and 59% (5140 CI 4260-6280) to be male.

# DISCUSSION

This is the first time a CHB prevalence of 0.2% for all Scotland has been predicted based on a combination of laboratory data and population demographics. This prevalence was slightly lower than the generally accepted figure of 0.3% for the UK population [2] and less than half of the most recent published prevalence of 0.45% in England and Wales [4]. These differences in predictions could be explained by differences in the ethnic composition of the Scottish population (98% 'White' in the 2001 census) compared to the English population (91% 'White' in the 2001 census). Additionally, these differences may be explained by different assumptions made about the prevalence of CHB in men compared to women. In the study with 0.3% CHB [2] it was assumed that compared to women, men had a two-fold higher risk, while our study indicated that gender differences in CHB prevalence were modified by ethnicity.

**Different CHB rates by ethnicity are likely caused by different modes of transmission — gender-neutral vertical transmission in the Chinese population and male-dominated horizontal transmission in the British & Other group. However, the increased risk for South Asian males compared to females of the same ethnicity were unexpected. Additional information on exposure risks through enhanced surveillance with local NHS Boards — responsible for following up new cases and contact tracing — is needed to explain these results.**

Given the high prevalence of CHB in ethnic minority groups, estimates of the ethnic composition of the general population are crucial but rarely available. In this study we present a novel approach to estimate the proportion of East Asians and South Asians in the Scottish population using a name-based ethnicity classification software applied to data on all people registered with a GP in Scotland (CHI database). We aim to extend this to other ethnic groups (in particular East Europeans and Africans) once data from the 2011 Scottish census become available. **Indeed, once detailed predictions of the number**

**of people living in England, Wales and Northern Ireland in 2011 by age group, sex and ethnicity become available, the number of people with CHB could be calculated for the whole of the UK.**

Approximately 3200 CHB infected people aged 15 and over with East Asian ethnicity were predicted to be living in Scotland in 2009, almost 40% of the total predicted CHB-positive population. The high number of infected people was due to the high prevalence of CHB in this population (e.g., 10% of women having had an antenatal test), while population estimates indicated that the proportion of adult East Asians in the Scottish population was low at 0.7%. Estimates of the large proportion of East Asians in the CHB infected population is supported by results from an (unpublished) **Scotland-wide** survey of Hepatitis B specialist services during 2009 which showed that approximately 50% of CHB patients at clinics were of Chinese ethnicity (G. Hawkings, personal communication).

Efforts to target the Chinese population and other ethnic minority groups for HBV testing and immunisation are already under way in Scotland (Andrew Bathgate, personal communication), the rest of the UK [21] and in other European Countries [22], [23]. Without such initiatives to diagnose CHB infected individuals (particularly men, as they are not covered through antenatal screening) cases go undiagnosed until symptoms develop.

Our study had several limitations. Whilst we give 95% confidence intervals for our prevalence estimates, we could only take account of the statistical uncertainty in the estimated prevalence of i) CHB in the antenatal population, ii) the relative risk of testing HBsAg-positive among women older than 44 years compared to those aged 15-29 years and men compared to women; and iii) the predictive values for Onomap classification.

In order to estimate the prevalence of CHB in men and in women older than 44 years, it was assumed that the relative risks, as calculated from HBV tests conducted in primary and secondary health care settings, were representative of differences in the general population. This assumption is supported by a comparison of CHB prevalences between women aged 15 to 29 years and those

aged 30 to 44 years in antenatal tests and in primary and secondary care. In antenatal tests of British and Other women, those aged 15 to 29 had a 1.6 times higher prevalence compared to those aged 30 to 44 years. In tests taken in primary and secondary care the rate between the two age groups was a similar 1.3. In antenatal tests of South Asian women, those aged 15 to 29 had a 0.5 times lower prevalence compared to those aged 30 to 44 years. In tests taken in primary and secondary care the rate between the two age groups was a similar 0.8. However, it is possible that the estimated relative risks are confounded by different indications for testing in the different age and gender strata. Reasons for HBV testing are not routinely recorded with the laboratory surveillance data and could therefore not be included in the analysis.

We relied on estimates for the PPV and NPV from Lakha *et al.* [16] to adjust for misclassification of Onomap in generating estimates of Scotland's population by ethnicity in 2009. The distribution of the Scottish population derived through this process was similar to that reported by Wohland *et al.* [20]. Slightly higher estimates of the proportion of South Asian males in the latter study would have little effect on the predicted prevalence of CHB in Scotland, a reduction by 0.008%.

In our analysis, people from low HBV prevalence countries (e.g. the UK and Northern Europe) were grouped together with people from medium and high prevalence countries (e.g. Eastern and Southern Europe, Africa) due to the sub-optimal performance of Onomap to accurately classify names relating to the latter ethnic groups [16]. The prevalence observed from antenatal tests in the British and Other stratum (0.13%) was therefore higher than in British women alone (0.06%) **; the observed higher prevalence in those aged 15-29 compared to 30-44 years in the British and Other ethnicity could be due to a higher proportion of the younger group originating from medium and high prevalence countries.**. Further work is therefore needed to accurately predict the numbers of people with CHB within ethnic minority groups, other than East and South Asians.

16

Name-based 'ethnicity' classification is routinely used by Public Health England to assign ethnicity to laboratory test data [24]. In the absence of information on ethnic background of people tested for CHB in various settings, use of full names was here the only way to attribute ethnicity and calculate reliable estimates of national prevalence. Consent to generate ethnicity data from the thousands of individuals undergoing a HBV test was not retrospectively sought. Permission to use personal information was however granted by HPS' Clinical Governance Committee, as the public health gain of this exercise far outweighed the negligible risk to the individual. Name-based 'ethnicity' classification was not available for tests conducted in sexual health clinics or occupational health in Scotland, because these usually do not share personal identifiers without explicit permission.

In conclusion, this study has estimated that in Scotland in 2009 approximately 8700 people over the age of 14 (0.2% of this population) were infected with CHB. Slightly less than half of the infected population belongs to ethnic minority groups. There is a need to identify the undiagnosed population and thereby to prevent disease progression into liver cirrhosis and liver cancer and to prevent spread of disease into the non-infected population. To this end, novel strategies should be developed to promote and offer testing to people at increased risk of CHB. **Indeed, the Scottish Government's Sexual Health and Blood-borne Virus Framework aims to better targeted testing, early diagnosis and effective treatment and care of blood borne virus infections [25]. Our study has also demonstrated that even without access to population sizes from census statistics, the size of the infected population in a nation can be estimated.**

# REFERENCES

## References

[1] World Health Organization. Fact sheet No204 [homepage on the Internet]. Word Health Organization; 2012. Available from: `http://who.int/mediacentre/factsheets/fs204/en/index.html`.

[2] Department of Health. Getting ahead of the curve: a strategy for combating infectious diseases (including other aspects of health protection). A report by the chief medical officer. Department of Health; 2002.

[3] Boxall EH, Flewett TH. Prevalence of HBsAg in UK population. British Medical Journal. 1987;294(6566):57.

[4] Hahné S, Ramsay M, Balogun K, Edmunds WJ, Mortimer P. Incidence and routes of transmission of hepatitis B virus in England and Wales, 1995-2000: implications for immunisation policy. Journal of clinical virology. 2004 Apr;29(4):211–220.

[5] Gay NJ, Hesketh LM, Osborne KP, Farrington CP, Morgan-Capner P, Miller E. The prevalence of hepatitis B infection in adults in England and Wales. Epidemiology and Infection. 1999 Feb;122(1):133–138.

[6] Boxall E, Skidmore S, Evans C, Nightingale S. The prevalence of hepatitis B and C in an antenatal population of various ethnic origins. Epidemiology and Infection. 1994 Dec;113(3):523–528.

[7] Caley M, Fowler T, Greatrex S, Wood A. Differences in hepatitis B infection rate between ethnic groups in antenatal women in Birmingham, United Kingdom, May 2004 to December 2008. Euro surveillance : bulletin europeen sur les maladies transmissibles = European communicable disease bulletin. 2012 Jul 26;17(30):20228.

[8] Siddiqui MR, Gay N, Edmunds WJ, Ramsay M. Economic evaluation of infant and adolescent hepatitis B vaccination in the UK. Vaccine. 2011 Jan 10;29(3):466–475.

[9] Salisbury D, Ramsay M, Noakes K. Immunisation against infectious disease. London: The Stationary Office; 2006.

[10] Buti M, Brosa M, Casado MA, Rueda M, Esteban R. Modeling the cost-effectiveness of different oral antiviral therapies in patients with chronic hepatitis B. Journal of hepatology. 2009 Oct;51(4):640–646.

[11] Custer B, Sullivan S, Hazlet T, Iloeje U, Veenstra D, Kowdley K. Global epidemiology of hepatitis B virus. Journal of Clincal Gastroenterology. 2004;38(10):158–168.

[12] Veldhuijzen IK, Toy M, Hahné SJ, Wit GAD, Schalm SW, de Man RA, et al. Screening and early treatment of migrants for chronic hepatitis B virus infection is cost-effective. Gastroenterology. 2010 Feb;138(2):522–530.

[13] Wong WW, Woo G, Heathcote EJ, Krahn M. Cost effectiveness of screening immigrants for hepatitis B. Liver international : official journal of the International Association for the Study of the Liver. 2011 Sep;31(8):1179–1190.

[14] Department of Health. Screening of pregnant women for hepatitis B and immunisation of babies at risk; 1998.

[15] National Institute for Health and Clincal Excellence. PH43 Hepatitis B and C - ways to promote and offer testing: guidance. NICE public health guidance 43. 2012;Available from: `http://www.nice.org.uk/nicemedia/live/14003/61863/61863.pdf`.

[16] Lakha F, Gorman DR, Mateos P. Name analysis to classify populations by ethnicity in public health: Validation of Onomap in Scotland. Public health. 2011 10;125(10):688–696.

19

[17] Tehami N, Mills P, Cameron S, Penrice G, Hague R, Mathers A. Outcome of the management of hepatitis B infection in pregnancy. Gut. 2011 April 01;60(Suppl 1):A248–A249.

[18] General Register Office for Scotland (GROS). Mid-2009 Population Estimates Scotland [database on the Internet]. General Register Office for Scotland. 2010. Available from: `http://www.gro-scotland.gov.uk/files2/stats/population-estimates/mid-2009/mid-2009-pop-est-scotland.pdf`.

[19] General Register Office for Scotland (GROS). Ethnicity and Religion Tables National [database on the Internet]. General Register Office for Scotland. 2011. Available from: `http://www.scrol.gov.uk/scrol/warehouse/NewWards_ER_N.jsp`.

[20] Wohland P, Burkitt M, Norman P, Rees P, Boden P, Durham H. ETH-POP Database, ESRC Follow on Fund 'Ethnic group population trends' [homepage on the Internet]. University of Leeds; 2012. Available from: `http://www.ethpop.org`.

[21] Larcher VF, Bourne J, Aitken C, Jeffries D, Hodes D. Overcoming barriers to hepatitis B immunisation by a dedicated hepatitis B immunisation service. Archives of Disease in Childhood. 2001 Feb;84(2):114–119.

[22] Cowan SA. Denmark scales up hepatitis B screening and vaccination for risk groups. Eurosurveillance. 2005 Nov 3;10(11):E051103.4.

[23] Veldhuijzen IK, Smits LJ, van de Laar MJ. The importance of imported infections in maintaining hepatitis B in The Netherlands. Epidemiology and Infection. 2005 Feb;133(1):113–119.

[24] Gungabissoon U, Balogun MA, Ramsay ME. Hepatitis C virus: laboratory surveillance in England and Wales, 1992-2004. Epidemiology and Infection. 2007 May;135(4):541–548.

[25] The Scottish Government. The Sexual Health and Blood Borne Virus Framework; 2011-15. The Scottish Government; 2011.

Table 4: Stratum specific predicted prevalence (%) and number (N) of HBsAg-positive adults in Scotland, 2009.

| Age group | (a) CHB Prevalence (%) (95%CI) | | (b) Prevalent number with CHB (N) (95%CI)[a][b] | | |
| | Females | Males | Females | Males | Total |
| --- | --- | --- | --- | --- | --- |
| British and other | | | | | |
| 15-29 | 0.16 (0.13-0.19) | 0.27 (0.20-0.35) | 780 (630-930) | 1350 (1020-1760) | 2130 (1670-2650) |
| 30-44 | 0.10 (0.08-0.13) | 0.17 (0.12-0.23) | 530 (390-670) | 830 (580-1120) | 1350 (990-1770) |
| >44 | 0.05 (0.04-0.07) | 0.08 (0.06-0.12) | 590 (430-800) | 850 (590-1220) | 1440 (1040-2000) |
| All ages | | | 1890 (1580-2250) | 3030 (2330-3910) | 4920 (3970-6050) |
| South Asian | | | | | |
| 15-29 | 0.43 (0.24-0.70) | 0.78 (0.36-1.43) | 50 (20-70) | 120 (60-230) | 170 (80-290) |
| 30-44 | 0.85 (0.50-1.34) | 1.54 (0.74-2.78) | 80 (50-130) | 230 (110-410) | 310 (160-530) |
| >44 | 0.54 (0.24-1.06) | 0.97 (0.38-2.16) | 30 (20-70) | 80 (30-180) | 120 (50-250) |
| All ages | | | 160 (110-230) | 430 (240-750) | 600 (360-960) |
| East Asian | | | | | |
| 15-44 | 10.27 (8.80-11.91) | 11.64 (8.62-15.60) | 1240 (980-1510) | 1330 (950-1840) | 2570 (1990-3280) |
| >44 | 9.31 (6.52-13.00) | 10.55 (6.85-16.04) | 290 (200-420) | 350 (220-540) | 640 (420-940) |
| All ages | | | 1520 (1200-1880) | 1680 (1190-2330) | 3200 (2460-4120) |
| All ethnicities | | | | | |
| 15-44 | | | 2670 (2340-3020) | 3860 (3210-4670) | 6530 (5630-7580) |
| >44 | | | 910 (730-1160) | 1280 (980-1710) | 2200 (1720-2850) |
| All ages | | | 3580 (3120-4080) | 5140 (4260-6280) | 8720 (7490-10230) |

[a]Rounded to nearest 10.

[b]Confidence intervals generated through bootstrapping.

*Epidemiology and Infection*

Use of laboratory-based surveillance data to estimate the number of people chronically infected with hepatitis B living in Scotland

C. Schnier, L. Wallace, K. Tempelton, C. Aitken, R.N. Gunson, P. Molyneaux, P. McIntyre, C. Povey, D. Goldberg, S. Hutchinson

Supplementary Material

**Calculation of the size of the Scottish population by ethnicity, sex and age group (Table 3) — adjustments for the predictive values of name classification using Onomap and for the predicted size of the population.** The positive predictive value of a test can be calculated as

$$PPV = \frac{a}{a+b}$$

where $a$ is the number of true positives and $a+b$ is the number of test positives. Similarly, the negative predictive value can be calculated as

$$NPV = \frac{c}{c+d}$$

where $c$ is the number of true negatives and $c+d$ is the number of test negatives. To estimate the adjusted total number of people in the CHI database that belong to ethnic group $i$, we calculated

$$P_i = PPV * (a+b) + (1 - NPV) * (c+d)$$

where $P_i$ denotes the adjusted number, $PPV * (a+b)$ denotes the number of people that were rightly classified by Onomap to belong to ethnic group $i$ and $(1 - NPV) * (c+d)$ denotes the number of people that were wrongly classified by the software as not belonging to ethnic group $i$.

Once $P_i$ had been calculated, the size of the different age and sex strata in the

extract from CHI could be estimated assuming that both age and sex had been recorded in the CHI database without misclassification and that misclassification by Onomap was not modified by age group or sex.

As the number of people in the CHI database by age and sex does not accurately estimate the size of the Scottish population, estimates for the latter were extracted from GRO. People in the different age and sex strata in the Scottish population were then assigned an ethnic group based on the (Onomap-adjusted) distribution of ethnic groups in the CHI extract.