Prevalence and risk factors of hepatitis B virus infection in Middle-Limburg Belgium, year 2017: Importance of migration.

Peer-reviewed author version

KOC, Ozgur; KREMER, Cécile; BIELEN, Rob; Buscchots, D.; HENS, Niel; Nevens, F. & ROBAEYS, Geert (2019) Prevalence and risk factors of hepatitis B virus infection in Middle-Limburg Belgium, year 2017: Importance of migration.. In: Journal of medical virology, 91 (8), p. 1479-1488.

DOI: 10.1002/jmv.25457
Handle: http://hdl.handle.net/1942/28454

Prevalence and risk factors of hepatitis B virus infection in Middle-Limburg Belgium, year 2017: importance of migration.

Özgür M Koc[1-3,*], Cécile Kremer[4], Rob Bielen[1,2,], Dana Buscchots[1,2], Niel Hens[4,5], Frederik Nevens[6], Geert Robaeys[1,2,6]

1. Department of Gastro-enterology and Hepatology, Ziekenhuis Oost-Limburg, Schiepse Bos 6, 3600 Genk, Belgium.
2. Faculty of Medicine and Life Sciences, Hasselt University, Martelarenlaan 42, 3500 Hasselt, Belgium.
3. Department of Medical Microbiology, School of Nutrition and Translational Research in Metabolism, Maastricht University Medical Centre, P.O. Box 5800, 6202 AZ, Maastricht, the Netherlands
4. Interuniversity Institute for Biostatistics and statistical Bioinformatics (I-Biostat), Hasselt University, Martelarenlaan 42, 3500 Hasselt, Belgium
5. Centre for Health Economic Research and Modelling Infectious Diseases, Vaccine and Infectious Disease Institute, University of Antwerp, Universiteitsplein 1, 2610 Antwerp, Belgium
6. Department of Gastro-enterology and Hepatology, KU Leuven, Herestraat 49, 3000 Leuven, Belgium.

Email: ozgur.koc@uhasselt.be, cecile.kremer@uhasselt.be, rob.bielen@uhasselt.be, dana.busschots@uhasselt.be, niel.hens@uhasselt.be, Frederik.nevens@uzleuven.be, geert.robaeys@zol.be

[*]**Corresponding author:**
Özgür M Koc, MD,
Department of Gastro-enterology and Hepatology, Ziekenhuis Oost-Limburg, Schiepse Bos 6, 3600 Genk, Belgium.
Tel: +32.89.321519, Fax: +32-89-327916, E-mail: ozgur.koc@uhasselt.be

Running title: HBV prevalence in Middle-Limburg Belgium

Length of paper: 3067 words
Total number of figures and tables: 5

Acknowledgements: This study is part of the 'Limburg Clinical Research Program' (LCRP), supported by the foundation Limburg Sterk Merk, province of Limburg, Flemish government, Hasselt University, Ziekenhuis Oost-Limburg and Jessa Hospital.

Author's contributions: ÖK, CK, RB, DB, NH, FN and GR contributed to the conception and design of the study. ÖK and RB were involved in planning and GR supervised the work. ÖK and RB collected the data. CK performed the analysis. ÖK drafted the first version of the paper, the co-authors critically revised the article and approved the final version, including the authorship list.

Access to data: The datasets used during the current study will be available from the corresponding author on reasonable request.

Conflicts of interest: ÖK has received a travel grant and research grants from Gilead. CK has no conflicts of interest. RB has received travel grants from Abbvie, MSD and Gilead to attend scientific congresses and research grants from Gilead and MSD. DB has received a travel grant from Abbvie. NH has no conflicts of interest. FN has received research grants, consultancy agreements and travel grants from UCB, Ipsen, Roche, Astellas, Ferring, Novartis, Janssen-Cilag, Abbvie, Gilead, CAF, Intercept, Gore, BMS, MSD, Promethera Biosciences, Ono Pharma, Durect. GR has received research grants from Abbvie, MSD, Janssen Pharmaceuticals, and has acted as a consultant/advisor for Abbvie, MSD, Gilead and BMS.

Funding/Sponsor: This study was supported by the Gilead Sciences by an unrestricted grant registered as V-2112 at Hasselt University.

Role of the Funder/Sponsor: The funding source had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Clinical trial number:
This clinical trial is registered at clinicaltrials.gov (NCT03425513).

Abstract

Background & Aims The hepatitis B virus (HBV) prevalence study performed in 2003 in Belgium is believed to be underestimating HBV prevalence due to underrepresentation of the non-Belgian population. Therefore, we assessed the prevalence and risk factors of HBV infection in a multi-ethnic region situated in Middle-Limburg Belgium, in 2017.

Methods Between May and November 2017, blood samples and questionnaires were taken from patients who presented at the emergency department of a large educational hospital. Blood samples were tested for hepatitis B surface antigen (HBsAg) and hepatitis B core antibodies (anti-HBc). A sample size of 1,000 persons was required to obtain a representative sample of the general Middle-Limburg population.

Results Of the 1,131 patients screened, the overall HBsAg prevalence was 0.97% with differences between Belgians (0.67%) and first-generation-migrants (2.55%), (p = .015). Five (45.5%) out of 11 HBsAg positive individuals were not aware of their HBV status. All five (100%) newly diagnosed HBsAg positive patients had further clinical evaluation and all had a normal level of alanine-aminotransferase. The prevalence of anti-HBc was 8.4%, and was significantly associated with age-gender-ethnicity interaction, presence of HBV infected household member, hepatitis C virus infection, men who have sex with men, and haemodialysis.

Conclusions In this area with large immigrant populations, we found a higher prevalence of HBV infection compared to the nationwide study of 2003. National HBV screening for first-generation migrants is needed as this high-risk group will go unnoticed due to the possible incorrect interpretation of normal alanine-aminotransferase values.

Length of abstract: 246 words

Key words: Hepatitis B; Prevalence; Risk factors; Migrants; Belgium

Introduction

Hepatitis B virus (HBV) remains a global health problem given that an estimated two billion people have been exposed to this virus and 257 million people live with chronic HBV infection worldwide in the year 2017.[1] Since the onset of HBV infection is generally asymptomatic, many of the HBV infected patients may not be aware of their infection status, thereby increasing the likelihood of infecting others.[2] Chronic HBV infection was also responsible for up to 887,000 deaths in 2015, making it together with hepatitis C virus (HCV) infection the 7th leading cause of mortality worldwide.[1,3]

The prevalence of chronic HBV varies upon the geographic distribution, from high-prevalence areas (≥8% hepatitis B surface antigen (HBsAg) positive), to intermediate-prevalence areas (2-7% HBsAg positive) and low-prevalence areas (<2% HBsAg positive).[4] Western Europe is among the low-prevalence areas, but subgroups of higher HBV prevalence may exist in regions with large immigrant populations. Even though migrants comprise only one in 20 European citizens, they account for one in four of all chronic HBV infections. In Belgium, 52% of chronic HBV cases are estimated to be among migrants from intermediate- and high-prevalence areas. This number is even >70% in countries such as Austria, Ireland, the Netherlands and United Kingdom.[5]

Both prevalence studies in Flanders (Belgium), one in 1993-1994 and one in 2003, showed that 0.7% of the population was HBsAg positive.[6,7] However, the prevalence of chronic HBV infection in the population-based study in 2003 is believed to be an underestimation due to underrepresentation of the non-Belgian population. Moreover, there are no data regarding the predominant risk factors for HBV infection in the Belgian general population. The recognized risk factors of HBV infection include having an HBsAg positive mother, HBsAg positive

close family member, HBsAg positive sex partner, multiple (unsafe) heterosexual contacts, men who have sex with men (MSM), intravenous drug use (IDU), blood or blood-product transfusion, haemodialysis, invasive healthcare procedure or dental treatment and tattooing or body piercing.[8]

Even in low-prevalence areas such as Belgium, HBV prevention and control is a public health priority, particularly since safe and effective vaccines are available. In addition to primary prevention, recent advancements in the treatment of chronic HBV infection now allow secondary prevention.[9,10]

Since chronic HBV infections are often asymptomatic, sero-epidemiological studies are needed (1) to understand the extent and importance of this public health problem, (2) to identify population subgroups with an increased prevalence of infection and (3) to foresee its future impact on the health system and ensure an adequate allocation of financial resources.

The aim of the present study was to evaluate the current prevalence and risk factors of HBV infection in a multi-ethnic region situated in Middle-Limburg Belgium.

Patients and methods

Study population and sample design

This epidemiological study was conducted in a multi-ethnic region of about 240,803 inhabitants in the region of Middle-Limburg between May and November 2017. All patients between the age of 18 and 70 who presented at the emergency department of a large educational hospital were eligible for this study. Upon written informed consent, a blood sample was taken and a face-to-face questionnaire, assessing demographics (age, gender,

158 country of birth, mother's country of birth), known viral hepatitis status (HBV positive, HCV
159 positive, HIV positive), and risk factors was performed. These risk factors included having an
160 HBsAg positive close family member, multiple (unsafe) sexual contacts, MSM, IDU, blood
161 or blood-product transfusion, haemodialysis, invasive healthcare procedure or dental
162 treatment (i.e. surgery), healthcare worker, cultural or ritual intervention, and tattooing or
163 body piercing. Immigrants who were not born in Belgium were considered first-generation
164 migrants (FGMs).
165
166 Trained personnel collected a blood sample for serological testing for each patient. The
167 Ziekenhuis Oost-Limburg laboratories tested all blood samples for HBsAg and hepatitis B
168 core antibodies (anti-HBc) using an electrochemiluminescence assay (Cobas 8000 e602,
169 Roche, Germany). The interpretation of positive and negative results was carried out as
170 recommended by the test producer.
171
172 The culturally targeted, multilingual (Dutch-, English- and Turkish-speaking) first author
173 attempted to contact persons with HBsAg positive results via telephone one week after their
174 screening. At least six attempts at three different times of the day were made. Persons not
175 reached within six months were considered lost to follow-up. The first author encouraged and
176 invited all newly diagnosed HBsAg positive patients to enter an Outpatient Hepatology
177 Department. A clinical work-up and treatment program was proposed for newly diagnosed
178 patients based on European Association for the Study of the Liver (EASL) guidelines.[11]
179
180 **Ethics approval and Trial registration**
181 The study was approved by the local Medical Ethical Committees (16/072U), and was
182 conducted in accordance with the provisions of the Declaration of Helsinki and its

amendments. Good clinical practice guidelines were followed throughout the study. The study is registered at clinicaltrials.gov (NCT03425513).

Statistical analyses

Sample size calculation was performed with the aid of Epi Info (Version 7.2.1.0, Atlanta, GA). The number of patients needed per ethnicity was calculated so that the results of the sample group agreed with those of the Middle-Limburg population with a confidence interval of 95%. The expected frequency for FGMs was 21% (*Limburg in cijfers*, personal communication). The sample size calculation suggested a total of 1,000 tested persons would need to be included.

Survey data were entered bed-side into a secure electronic database Castor EDC (Castor Electronic Data Capture, Ciwit Bv, Amsterdam, the Netherlands). Analysis of variance (ANOVA) was used to test for differences in mean age and prevalence between different ethnic groups.

For calculating the effect of universal HBV vaccination in Belgium, we stratified HBsAg and/or anti-HBc prevalence into two birth cohorts (born before 1987 and born after 1987). The cutoff was chosen as the vaccination program since September 1999 with catch-up vaccination for one age cohort with the age range of 10-13 years covered children born after 1987.[12]

In order to correct for differences between the sample and the Middle-Limburg population, weighted tests were conducted. Weighting was done based on the combination of age and ethnicity (Belgian vs. non-Belgian (i.e. FGM)).[13] Weighted chi squared tests were used to assess significant associations between the evaluated risk factors and HBsAg or anti-HBc prevalence. Risk factors that were shown to be significantly associated ($p \leq .10$) to HBsAg or anti-HBc in these univariate analyses were included in a weighted multiple logistic regression

model. In these models, Firth's bias adjustment was used to account for data sparseness.[14] Model reduction was done in a backward stepwise manner based on the .05 significance level. Several classification methods were also applied to the data on anti-HBc positivity for prediction purposes. First, a simple classification tree was constructed based on an almost completely balanced training dataset (i.e. 40% of the subjects were anti-HBc positive) to identify characteristics that explain the outcome in the best way. Bagging and random forests were then used in an attempt to improve the error rate of this classification tree. Since the predictions based on (almost) balanced data overestimated the true proportion of anti-HBc positives, a correction was made.[15] The classification models were then compared to the weighted logistic regression model in terms of prediction accuracy. Data analyses were performed using RStudio (Version 1.0.136, Boston, MA).

Results

Characteristics of the study population

Of the 1,537 individuals invited, 1,131 (73.6%) completed the questionnaire and donated blood. The study included 605 men (53.5%) and 526 women (46.5%), with a mean age of 46 years (95% CI 45.2-46.9 years). FGMs comprised 20.8% of the study population. FGMs were mainly born in the Netherlands (32.8%), Turkey (18.7%), and Italy (12.3%). Gender distribution was similar in both ethnic groups (Belgian and FGM, $p = .629$). When compared to the Belgian patients (mean age of 46 years, 95% CI 45.0-46.9 years), FGMs (mean age of 47 years, 95% CI 44.8-48.3 years) were not significantly older ($p = .536$).

Patients who refused participation had a mean age of 48 years (95% CI 43.3-46.3 years) and 52.1% were male. This did not significantly differ from patients willing to participate ($p = .135$ and $p = .699$ for age and gender, respectively). Reasons for not participating were: fear

of needles (14.5%), too sick (26.9%), other worries than testing (27.3%), does not want to know viral hepatitis status (8.1%), already participating in a lot of studies (0.3%), and other (22.9%).

Prevalence of anti-HBc antibodies

In two patients, blood sample volumes were insufficient for testing and in 22 patients, no anti-HBc blood sample was taken. Anti-HBc positivity was 8.4% among the remaining 1,107 participants. Anti-HBc prevalence in Belgian individuals born after 1987 was 1.6%. This number was higher for those born before 1987 (7.0%, $p = .001$). Current or past HBV infection in FGMs was apparent in 49/231 (21.2%) with differences in those born in the Netherlands (5/75, 6.7%), Turkey (16/44, 36.4%), Italy (8/28, 28.6%), other low endemic countries (2/30, 6.7%) and other intermediate or high endemic countries (18/54, 33.3%), $p < .001$. In the weighted chi squared tests, age-gender-ethnicity interaction ($p < .001$), not shown in table), living with an HBV infected person ($p < .001$), HCV infection ($p = .004$), MSM ($p = .001$), blood transfusion before 1972 ($p = .026$), haemodialysis ($p < .001$), cultural or ritual intervention ($p = .015$) and having a tattoo or body piercing ($p = .041$) were significantly associated with current or past HBV infection, i.e. anti-HBc positivity (Table 1).

Since only 2 out of 7 patients that have had a blood transfusion before 1972 were anti-HBc positive, this risk factor was not included in any further analyses. Parameter estimates of the weighted multiple logistic regression model are shown in Table 2. It can be seen that, compared to 18-39 years old Belgian females, there was a significantly increased risk of anti-HBc positivity in all FGMs, as well as in 40-70 years old Belgian females ($p = .039$). Young Belgian males had a lower risk of anti-HBc positivity, although not statistically significant. Furthermore, living with an HBV infected person ($p < .001$), reporting HCV infection ($p =$

.005), being on haemodialysis treatment ($p < .001$), and being MSM ($p = .002$) appeared to be associated with a higher risk of anti-HBc positivity.

Anti-HBc classification models

In growing the classification tree, only the risk factors that were shown to be significantly associated ($p < .10$) to anti-HBc prevalence (in the weighted chi squared tests) were used as input (except for having received a blood transfusion before 1972, for the same reason mentioned above). The obtained classification tree is shown in Figure 1. The predictive accuracy of this tree was 73.0% based on the ROC curve (see solid line in Figure 2), with a sensitivity of 57.1% and specificity of 80.1%. The predictive accuracy was also calculated for the weighted logistic regression model, giving a predictive accuracy of 77.5% with a sensitivity and specificity of 68.8% and 77.8%, respectively. From Figure 2 it can be concluded that the weighted logistic regression model using Firth's bias adjustment performed best in predicting anti-HBc prevalence in this study.

Prevalence of HBsAg

Of the 1,131 patients tested, 11 (1.0%) were HBsAg positive. Five (45.5%) HBsAg positive patients were born in Belgium, two (18.2%) in Italy, one (9.1%) in Turkey and three (27.2%) in other intermediate or high endemic countries (one in Saudi Arabia, Iran, and Kenya), $p = .136$. None of the HBsAg positive patients were younger than 30 years of age (born after 1987) compared to 11 HBsAg positive patients born before 1987, $p = .282$. The HBsAg prevalence in FGMs was 2.55% (6/235), with differences in those born in the Netherlands (0/77, 0.0%), Turkey (1/44, 2.27%), Italy (2/29, 6.9%), other low endemic countries (0/30, 0.0%) and other intermediate or high endemic countries (3/55, 5.5%), $p = .285$.

Table 3 shows the prevalence of HBsAg by different risk factors. In the weighted chi squared tests, HBsAg positivity was significantly associated with ethnicity (p = .015), having an HBV infected household member (p = .036), MSM (p = .016), and tattooing or body piercing (p = .077). Logistic regression was not conducted for HBsAg due to the low number of chronically infected patients.

Linkage to care

Of the 11 HBsAg positive individuals, five (45.5%) were not aware of their HBV status and the other six (54.5%) were already linked to care. Thus, the percentage of newly diagnosed HBsAg positive patients was 0.44% (5/1131). All five (100%) patients had further clinical evaluation which revealed that all had a normal level of alanine aminotransferase (ALT <40 U/L), all were hepatitis B e antigen (HBeAg) negative, all had HBV DNA levels below 2,000 IU/mL, all had fibrosis score F0-1 according to Metavir score, and none had evidence of liver cirrhosis on ultrasound. One of the five patients is currently being treated prophylactically in the Outpatient Hepatology Department before and during the administration of chemotherapy according to the EASL guidelines.

Discussion

This is the first study to assess the seroprevalence of HBV infection in a multi-ethnic Belgian region including a considerably large proportion of foreign-born individuals. Besides estimating the seroprevalence in a multi-ethnic region, risk factors associated to HBV infection were evaluated, which may assist physicians, public health practitioners and policymakers in eliminating hepatitis B as a public health threat by 2030.[16]

The principal findings of the present study can be summarized as follows. First, there was an overall HBsAg seroprevalence of 1.0%, with higher prevalence in FGMs (2.55%) compared to Belgians (0.7%). Second, none of the HBsAg positive patients were born after 1987. Third,

an anti-HBc prevalence of 8.4% was found in the multi-ethnic region situated in Middle-Limburg with age-gender-ethnicity interaction, having an HBV infected household member, reporting HCV infection, being MSM, and ever having been on haemodialysis treatment being significantly associated with past or current HBV infection. Fourth, this study demonstrated an excellent linkage to care with five out of 11 (45.5%) HBsAg positive individuals not being aware of their HBV status and further clinical evaluation showing normal levels of alanine-aminotransferase in all five patients.

In this study, about 1.0% of all patients appeared to be HBsAg positive and 8.4% showed evidence of HBV exposure. Until now, the overall prevalence in Belgium was estimated to be lower, i.e. 0.7% positive for HBsAg and 6.4% positive for anti-HBc.[6-7] This discrepancy can be explained by the fact that both of these previous epidemiological studies also included individuals aged 0-17 years. Moreover, selection bias can explain a possible underestimation in 2003 as the recruiting of participants in the general population by mail probably missed people from certain risk groups (e.g. migrants). In the present study, we included patients aged 18-70 years who presented at the emergency department of a large educational hospital in a multi-ethnic region situated in Middle-Limburg. Furthermore, the region of Middle-Limburg has a large immigrant population, with 20.8% of the study population being FGMs, in contrast to 8.0% in Flanders, Belgium. In this respect, we found a higher HBsAg prevalence in FGMs and especially in those born in intermediate or high endemic countries, highlighting the fact that migrants are an important risk group for chronic HBV infection.[17-18]

The HBsAg prevalence in the general population of the neighbouring countries ranges from 0.1% in Ireland to 0.8% in Spain.[19] In line with our findings, a higher HBV prevalence was found in FGMs born in intermediate- or high-prevalence areas in comparison to the native population of countries such as France, Germany, the Netherlands and Spain.[20-23]

The impact of implementation of a universal free-of-charge hepatitis B vaccination in Belgium was also apparent from the results of the current study. None of the HBsAg positive patients were born after year 1987. Anti-HBc prevalence was also lower in Belgian individuals born after 1987, when compared to those born before 1987. After all, universal infant hepatitis B vaccination with catch-up in adolescents aged 10-13 years began in September 1999 in Belgium.[12] Consequently, the vaccination program covered children born after 1987. To evaluate the effects of a universal HBV vaccination program in <20 year-olds, prevalence of seroprotection and HBV infection were assessed by Theeten et al.[24] They demonstrated that the prevalence of HBV infection remained low in Belgium and that overall high levels of 'vaccinated' serostatus were achieved in infants as well as in adolescents.

We also analysed the risk factors associated to anti-HBc positivity. The most prominent risk factors were being FGM and age between 40 and 70 years (except for Belgian males), having an HBV infected household member, reporting HCV infection, being MSM, and ever having received haemodialysis, with 22.22%, 26.32%, 26.67%, 33.33% and 36.36% testing positive for anti-HBc, respectively. A comparison of different classification methods revealed that the weighted logistic regression model performed best in classifying patients as either anti-HBc positive or negative, although the identified risk factors in this model were the same as those found in a classification tree analysis.

The implementation of a culturally and linguistically appropriate health care provider in the present study could explain the high linkage to care of newly diagnosed HBsAg positive patients.[25] Moreover, all newly diagnosed HBsAg positive patients underwent ALT determination and normal ALT levels were found in all of them. The Belgian nationwide epidemiological data support the findings of our study and emphasized that 80% of the newly diagnosed chronic HBV patients had normal ALT levels at diagnosis.[26] A major limitation of ALT as a biochemical marker of liver disease is that its levels often fluctuate over time during

the variable course of chronic HBV infection and may fail to identify patients with necro-inflammatory activity or fibrosis.[27]

There are some limitations to the present study. First, by screening at the emergency department, a bias towards subjects in certain risk groups may occur. However, these risk groups might have been underrepresented in the previous population-based study in Belgium.[7] Inclusion of certain risk groups in our study also allowed us to determine the most prominent risk factors for HBV prevalence in Belgium for the first time. Second, due to logistical factors (e.g. limited daily enrolment time, limited study research team) not every eligible participant could be informed about the study. Third, a concern is that, even though the patients' demographics, known viral hepatitis status and risk factors were recorded, certain risk behaviours could have been underreported as the questionnaire was performed using a face-to-face interview. Inferences from this study should also be drawn with caution since it is difficult to establish causal pathways from cross-sectional studies as our present study. Thus, we have attempted only to identify risk factors associated with HBV infection using odds ratios. This study could also be underpowered to find significant associations between certain risk factors and HBsAg or anti-HBc, as the study was powered in such a way that the distribution per ethnicity was similar to that of the Middle-Limburg population, and not to specific risk factors. However, good predictive accuracy was shown for the weighted multiple logistic regression model, providing evidence for the significant impact of abovementioned risk factors for anti-HBc.

In conclusion, this study shows that the HBsAg and anti-HBc seroprevalences in a multi-ethnic region in Middle-Limburg are higher than those previously found nationwide, probably because high-risk groups such as FGMs are more present. Since all newly diagnosed HBsAg positive patients had normal levels of ALT, national HBV screening for individuals born in intermediate or high endemic countries is needed as this high-risk group will go unnoticed

due to the possible incorrect interpretation of normal ALT values. In order to adapt or to adopt screening practices and preventive measures, sero-epidemiological studies should not only be done nationwide, but also locally in multi-ethnic regions.

References

1. World Health Organization. Hepatitis B. Geneva, Switzerland: World Health Organization; 2018. Available at: http://www.who.int/mediacentre/factsheets/fs204/en/. Accessed 21 January 2019.

2. Lin SY, Chang ET, So SK. Why we should routinely screen Asian American adults for hepatitis B: a cross-sectional study of Asians in California. Hepatology (Baltimore, Md). 2007;46(4):1034-40.

3. Stanaway JD, Flaxman AD, Naghavi M, Fitzmaurice C, Vos T, Abubakar I, et al. The global burden of viral hepatitis from 1990 to 2013: findings from the Global Burden of Disease Study 2013. Lancet (London, England). 2016;388(10049):1081-8.

4. Mahoney FJ. Update on diagnosis, management, and prevention of hepatitis B virus infection. Clinical microbiology reviews. 1999;12(2):351-66.

5. Ahmad AA, Falla AM, Duffell E, Noori T, Bechini A, Reintjes R, et al. Estimating the scale of chronic hepatitis B virus infection among migrants in EU/EEA countries. BMC infectious diseases. 2018;18(1):34.

6. Beutels M, Van Damme P, Aelvoet W, Desmyter J, Dondeyne F, Goilav C, et al. Prevalence of hepatitis A, B and C in the Flemish population. European journal of epidemiology. 1997;13(3):275-80.

7. Quoilin S, Hutse V, Vandenberghe H, Claeys F, Verhaegen E, De Cock L, et al. A population-based prevalence study of hepatitis A, B and C virus using oral fluid in Flanders, Belgium. European journal of epidemiology. 2007;22(3):195-202.

8. Chu CJ, Keeffe EB, Han SH, Perrillo RP, Min AD, Soldevila-Pico C, et al. Hepatitis B virus genotypes in the United States: results of a nationwide study. Gastroenterology. 2003;125(2):444-51.

9. Janssen HL, van Zonneveld M, Senturk H, Zeuzem S, Akarca US, Cakaloglu Y, et al. Pegylated interferon alfa-2b alone or in combination with lamivudine for HBeAg-positive chronic hepatitis B: a randomised trial. Lancet (London, England). 2005;365(9454):123-9.

10. Lok AS, McMahon BJ, Brown RS, Jr., Wong JB, Ahmed AT, Farah W, et al. Antiviral therapy for chronic hepatitis B viral infection in adults: A systematic review and meta-analysis. Hepatology (Baltimore, Md). 2016;63(1):284-306.

11. EASL 2017 Clinical Practice Guidelines on the management of hepatitis B virus infection. Journal of hepatology. 2017.

12. Erwin Van Kerschaver (2008) The Belgian Expanded Programme on Immunization (EPI) and hepatitis B vaccination, Southern African Journal of Epidemiology and Infection, 23:1, 40-44.

13. Centraal Bureau voor de Statistiek. Bevolking; leeftijd, migratieachtergrond, geslacht en regio. Den Haag, the Netherlands: Centraal Bureau voor de Statistiek; 2018. Available at: http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=37713&D1=1-2,7-8&D2=0&D3=1,3-4&D4=16&D5=l&HDR=G4,T,G1,G2&STB=G3&VW=T (accessed 22/05/18).

14. Firth D. Bias reduction of maximum likelihood estimates. Biometrika. 1993;80(1):27-38.

15. Lemmens A, Croux C. Bagging and boosting classification trees to predict churn. Journal of Marketing Research. 2006;43(2):276-86.

16. World Health Organization. Combating hepatitis B and C to reach elimination by 2030. Geneva, Switzerland: World Health Organization; 2016. Available at: http://apps.who.int/iris/bitstream/10665/206453/1/WHO_HIV_2016.04_eng.pdf?ua=1 (accessed 22/05/18).

17. Rossi C, Shrier I, Marshall L, Cnossen S, Schwartzman K, Klein MB, et al. Seroprevalence of chronic hepatitis B virus infection and prior immunity in immigrants and refugees: a systematic review and meta-analysis. PloS one. 2012;7(9):e44611.

18. Kowdley KV, Wang CC, Welch S, Roberts H, Brosgart CL. Prevalence of chronic hepatitis B among foreign-born persons living in the United States by country of origin. Hepatology (Baltimore, Md). 2012;56(2):422-33.

19. European Centre for Disease Prevention and Control. Systematic review on hepatitis B and C prevalence in the EU/EEA. Stockholm: ECDC; 2016

20. Meffre C, Le Strat Y, Delarocque-Astagneau E, Dubois F, Antona D, Lemasson JM, et al. Prevalence of hepatitis B and hepatitis C virus infections in France in 2004: social factors are important predictors after adjusting for known risk factors. Journal of medical virology. 2010;82(4):546-55.

21. Wolffram I, Petroff D, Batz O, Jedrysiak K, Kramer J, Tenckhoff H, et al. Prevalence of elevated ALT values, HBsAg, and anti-HCV in the primary care setting and evaluation of guideline defined hepatitis risk scenarios. Journal of hepatology. 2015;62(6):1256-64.

22. Hahne SJ, De Melker HE, Kretzschmar M, Mollema L, Van Der Klis FR, Van Der Sande MA, et al. Prevalence of hepatitis B virus infection in The Netherlands in 1996 and 2007. Epidemiology and infection. 2012;140(8):1469-80.

23. Pedraza-Flechas AM, Garcia-Comas L, Ordobas-Gavin M, Sanz-Moreno JC, Ramos-Blazquez B, Astray-Mochales J, et al. Hepatitis B virus infection and vaccine-induced immunity in Madrid (Spain). Gaceta sanitaria / SESPAS. 2014;28(6):492-5.

24. Theeten H, Hutse V, Hoppenbrouwers K, Beutels P, Van Damme P. Universal hepatitis B vaccination in Belgium: impact on serological markers 3 and 7 years after implementation. Epidemiology and infection. 2014;142(2):251-61.

21. Perumalswami PV, Factor SH, Kapelusznik L, Friedman SL, Pan CQ, Chang C, et al. Hepatitis Outreach Network: a practical strategy for hepatitis screening with linkage to care in foreign-born communities. Journal of hepatology. 2013;58(5):890-7.

22. De Vroey B, Moreno C, Laleman W, van Gossum M, Colle I, de Galocsy C, et al. Hepatitis B virus and hepatitis C virus infections in Belgium: similarities and differences in epidemics and initial management. European journal of gastroenterology & hepatology. 2013;25(5):613-9.

23. Papatheodoridis GV, Manolakopoulos S, Liaw YF, Lok A. Follow-up and indications for liver biopsy in HBeAg-negative chronic hepatitis B virus infection with persistently normal ALT: a systematic review. Journal of hepatology. 2012;57(1):196-202.

Fig 1 Classification tree based on (almost) balanced training sample for anti-HBc

A ‘0’ indicates anti-HBc negative, while a ‘1’ indicates anti-HBc positive.

Abbreviation: anti-HBc: hepatitis B core antibodies; FGM: first-generation migrants, i.e. foreign-born persons; HCV: hepatitis C virus; HBV: hepatitis B virus; MSM: men who have sex with men.

519 **Fig 2 Comparison of different classification models for anti-HBc using ROC curves for**
520 **which the AUC is used to quantify the predictive accuracy**

521 In an attempt to reduce the tree misclassification error, bagging was applied. All risk factors
522 were included and B=400 trees were grown. Class weights were used to correct for the
523 (almost) balanced training sample. The predictive accuracy of this model however was only
524 55.7% (see dashed line). For the random forest model, at each iteration 12 randomly sampled
525 risk factors were included and B=5000 trees were grown. Class weights were again used. The
526 predictive accuracy of this model was slightly higher than for bagging, i.e. 55.9% (see dotted
527 line). The six most important variables from the random forest (RF) analysis were included in
528 a logistic regression model (see grey dotted line), which had a predictive accuracy of 73.6%,
529 still lower than for the abovementioned weighted logistic regression model (77.5%, see dot-
530 dashed line).

531 Abbreviation: anti-HBc: hepatitis B core antibodies; ROC: receiver operating characteristic; AUC: area under the ROC curve; RF: random
532 forest.

533