Global epidemiology of hepatitis B virus infection: New estimates of age-specific HBsAg seroprevalence and endemicity

J.J. Ott a, G.A. Stevens a, J. Groeger b, S.T. Wiersma a,∗

a World Health Organization, 20, Avenue Appia, 1211 Geneva 27, Switzerland
b Centers for Disease Control and Prevention, Atlanta, GA, USA

ARTICLE INFO

Article history:
Received 1 September 2011
Received in revised form 9 December 2011
Accepted 23 December 2011
Available online 24 January 2012

Keywords:
Hepatitis B
Hepatitis B surface antigen (HBsAg)
Epidemiology
Prevalence
Systematic review
Age distribution
Seroepidemiologic studies
Modeling

ABSTRACT

Objectives: Chronic hepatitis B virus infection is one of the most serious infections and a major
risk factor for deaths from cirrhosis and liver cancer. We estimate age-, sex- and region-specific
prevalence of chronic HBV infection and calculate the absolute number of persons being chronically
infected.

Methods: A systematic review of the literature for studies reporting HBV infection was conducted
and worldwide HBsAg seroprevalence data was collected over a 27-year period (1980–2007). Based on
observed data, age-specific prevalence and endemicity were estimated on a global level and for all world
regions for 1990 and 2005 using an empirical Bayesian hierarchical model.

Findings: From 1990 to 2005, the prevalence of chronic HBV infection decreased in most regions. This was
particularly evident in Central sub-Saharan Africa, Tropical and Central Latin America, South East Asia
and Central Europe. Despite this decrease in prevalence, the absolute number of HBsAg-positive persons
increased from 223 million in 1990 to 240 million in 2005. Age-specific prevalence varied by geographical
region with highest endemicity levels in sub-Saharan Africa and prevalence below 2% in regions such as
Tropical and Central Latin America, North America and Western Europe. Asian regions showed distinct
prevalence patterns with lower intermediate prevalence in South Asia, but up to 8.6% HBsAg prevalence
in East Asia. Strong declines were seen in South East Asian children.

Conclusion: Declines in HBV infection prevalence may be related to expanded immunization. The increasing
overall number of individuals being chronically infected with HBV, and the widespread global differences in
HBV prevalence call for targeted approaches to tackle HBV-related mortality and morbidity. HBV infection
prevalence data are needed at country and sub-national level to estimate disease burden and guide health and vaccine policy.

© 2012 Elsevier Ltd. Open access under CC BY-NC-ND license.

1. Introduction

Knowledge of region- and age-specific prevalence of hepatitis B infection is important for evaluating vaccination programs and national disease prevention and control efforts. Furthermore, any modeling and assessment of the disease burden associated with the hepatitis B virus (HBV) requires prevalence estimates. So far, global studies on HBV seroprevalence are limited and comprehensive data are not available for many countries. In addition, demographic changes and expanded vaccination can create new epidemiological patterns of the virus which impact on region-specific endemicity levels.

HBV is spread predominantly by percutaneous or mucosal exposure to infected blood and other body fluids with numerous forms of human transmission. The sequelae of HBV infection include acute and chronic infection, cirrhosis of the liver and primary liver cancer. The likelihood of progression to chronic infection is inversely related to age at the time of infection. Around 90% of infants infected perinatally become chronic carriers, unless vaccinated at birth. The risk for chronic HBV infection decreases to 30% of children infected between ages 1 and 4 years and to less than 5% of persons infected as adults [1,2].

Chronic HBV infection progresses nonlinearly through 3–4 phases, from the immune-tolerant phase to immune clearance or immununactive phase, to nonreplicative inactive phase and possible reactivation [3,4]. After infection with HBV, most patients either develop immunity (87–90%) and clear the infection or become chronic carriers. A lower percentage will develop liver disease or chronic active hepatitis with an increased risk of developing cirrhosis, liver cancer or both [5]. The fatality of these diseases as well as their attribution to hepatitis infection is well known: 600,000 HBV-related deaths were estimated to occur annually [6] and 73%
of all liver cancer deaths worldwide are due to hepatitis viruses, with much higher proportions in low and middle income countries [7].

The complex serology and natural history associated with HBV infection creates challenges for the assessment of HBV prevalence and the provision of comparable global estimates. This is due to the availability of multiple laboratory markers for hepatitis B infection. Antibodies and antigens associated with this infection include hepatitis B surface antigen (HBsAg), antibody to hepatitis surface antigen (anti-HBs), antibody to hepatitis B core antigen (anti-HBc), and IgM antibody subclass of anti-HBc (IgM anti-HBc). Some studies also report markers of high HBV replication such as hepatitis B e" antigen (HBeAg), antibody to HBeAg (anti-HBe), and quantitative HBV-DNA.

HBsAg is the main clinical marker indicating acute or chronic infection and prevalence as well as endemicity of HBV infection is defined by the presence of HBsAg [8]. HBsAg testing is the primary way to identify persons with chronic HBV infection and several characteristics of this serological marker increase the precision of HBsAg estimates, including high specificity, low serum persistence, low possibility of chronic cases losing HBsAg [3,8,9]. However, routine population surveillance of chronic viral hepatitis is currently rare. Standardized monitoring would help not only in quantifying the disease burden on a population level but also in determining the characteristics of infected individuals, avoiding further transmission and allocating appropriate treatment. This is particularly important for populous countries that have been previously categorized as highly endemic for chronic hepatitis B infection such as China, Indonesia, Nigeria and parts of Africa and Asia, where an immense absolute number of people live with the virus [6,8]. However, up to date region-specific and globally comparable chronic HBV prevalence data are lacking and no relevant meta-analysis has been published on this topic. In addition, the absolute number of individuals being chronically infected with HBV is not known.

In the light of this gap, the objective of our study is to estimate age- and region-specific HBsAg prevalence in 1990 and 2005 by conducting a systematic review and modeling HBsAg prevalence. This investigation is part of the Global Burden of Diseases, Injuries, and Risk Factors Study, which carries out complete systematic assessments of global data on disease and risk factors in order to produce comparable estimates for two time periods, 1990 and 2005.

We provide detailed HBsAg prevalence estimates that are categorized by the 21 regions as defined by the Global Burden of Disease Study and geographically mapped by endemicity level. The Global Burden of Disease regions are based on geographic regions or continents and are grouped based on child and adult mortality levels and major causes of death in each country (see Section 2).

2. Methods

2.1. Systematic review

A systematic review of published literature and data was conducted to identify articles reporting prevalence of hepatitis B, C, and D virus infection for all countries over a 27-year period, from 1980 to 2007. Articles published in this time frame were included regardless of when the data were collected. Within each database, medical subject heading terms and freetext search terms were included to identify article abstracts that contained (1) a term related to the hepatitis B, C or D virus or their markers of hepatitis infection and (2) a term related to either prevalence, incidence, or disease burden (search terms available by request from the author). Results were restricted to original research articles in English. Abstracts were screened and were required to report hepatitis B or C prevalence or incidence. Where abstracts were incomplete or missing, the full-text article was retrieved and reviewed manually to determine if it reported prevalence of hepatitis B or C virus infection. A total of 6064 English citations were found (3273 Medline, 2283 Embase, 508 Cinhal). Review articles, outbreak investigations and national infections disease notification reports were excluded since they provide information on incident or acute cases. Data reported in the article had to be reasonably representative of the general population rather than conducted among a special high-risk group (i.e. injecting drug users, HIV-positive individuals) or a population that was selected based on a risk factor for viral hepatitis or a condition associated with hepatitis infection (Fig. 1, exclusion criteria).

After applying manual de-duplication and the exclusion criteria on the abstract, 1233 articles were obtained (references are listed in Web Annex 1). These were further screened for the specified exclusion criteria and for HBsAg as the marker of interest in the full text before country-specific prevalence information was extracted. For one country, the United States (US), a representative primary national data source was available and data were included only from articles reporting prevalence from the National Health and Nutrition Examination Survey (NHANES) [10,11].

Articles only reporting HCV marker (222) and those reporting summary or other markers of HBV (82) were excluded; 396 articles were determined to meet all eligibility criteria.

2.2. Analysis and modeling of data

Age-, sex- and country-specific HBsAg seroprevalence data were extracted. Additional information obtained was: primary author, year of publication, number of individuals tested for HBsAg, laboratory test/method, and study year. If the year was missing, two years prior to publication was assumed as the study year. In case age was not further specified, it was imputed based on contextual information such that pregnant women and those giving birth were assumed to be 15–49 years, army recruits and soldiers between 18 and 45 years, blood and organ donors between 17 and 65 years, and school children between 5 and 15 years. Extracted data were grouped according to 21 Global Burden of Disease Regions (Web Annex 2) and assigned a quality rating based on population size, sampling and representativeness of the general population. Using the extracted study seroprevalence data, prevalence of HBsAg was modeled using Dismod III v3.0, a generic disease modeling system [12]. DisMod III aims to synthesize data to generate estimates of the disease burden associated with more than 200 diseases and over 20 major risk factors for the Global Burden of Disease 2010 study [13]. DisMod III models multiple disease parameters, including incidence, prevalence, remission, and mortality, in order to ensure consistency among the parameters. Data on each of these parameters are synthesized using a hierarchical empirical Bayesian model to make estimates for 21 world regions based on observed data in each modeled region, data observed in other regions, and data from other time periods (by estimating a time trend). Briefly, Dismod III first fits an empirical prior estimate separately for each disease parameter (e.g., prevalence and incidence). The empirical prior has the following elements: geographic hierarchy, in which estimates for each region are informed by data from the same region and to a lesser extent data from other regions; a flexible age pattern; a linear time trend; and an offset for data on males. Second, for each period (1980–1996 and 1997–present), sex, and region (of 21 world regions), Dismod fits a Bayesian model using all data in that [1] The definition of chronic HBV infection and the respective marker used slightly varied by period of survey conduction. From 1988 on, HBV infection was defined as the presence of anti-HBc and HBsAg. We have included this estimate since all HBsAg positive individuals should be expected to have anti-HBc.
time–sex–region group and empirical priors for all epidemiological parameters, generating posterior estimates of incidence, prevalence, remission, and mortality that are internally consistent. In our model, the empirical priors for incidence, remission, and mortality were uninformative; thus the posterior was informed only by prevalence data. Like the empirical prior, the posterior models also incorporate linear time trends, flexible age patterns, and offsets for data on males. Applying region-specific population figures for 1990 and 2005, age- and region-specific HBsAg prevalence was used to calculate the absolute number of individuals chronically infected with HBV.

For the purpose of generating endemicity maps, endemicity levels of HBsAg were defined as low (<2%), lower intermediate (2–4%), higher intermediate (5–7%) and high (≥8%). HBsAg infection levels have traditionally been described according to three categories of endemicity indicating the proportion of the population being seropositive for HBsAg [8,14]. However, given that there are regions very close to low endemicity (<2%) (e.g. 2.0% prevalence among children in Western Europe and Central Latin America in 2005) and others rather close to high endemicity (≥8%) (e.g. 7.8% prevalence among adults in Eastern Asia in 2005) the split of the intermediate category better reflects regional differences and their implications.

3. Results

We identified 396 studies of HBsAg prevalence after applying all inclusion and exclusion criteria. To illustrate global endemicity, estimated and categorized HBsAg prevalence in 2005 is shown in Map 1 for children (5–9 years) and in Map 2 for adults (19–49 years) applying world population age weights.

The pattern of age-specific HBsAg prevalence varied greatly by region and the trend of a decreasing prevalence with age was more evident in 1990 as compared to 2005, where some regions, e.g. South East Asia showed an exceptional increase with age. For most regions, predominantly Tropical Latin America, West sub-Saharan Africa, Australasia, and North Africa, Figs. 2–5 indicate an overall decrease in HBsAg prevalence between 1990 and 2005. East Asia and Western Europe experienced some increase.

Global differences between males and females were small, although females had a lower overall HBsAg prevalence of 3.5% in

---

**Fig. 1. Flowchart of article screening.**

*Additional exclusion criteria applied in full text screening and not identified as such during first article review.*

(a) Populations of persons at high risk for hepatitis including those with diseases related to HBV such as acute (viral) hepatitis cases, liver cancer, and cirrhosis; HBV-positives.

(b) Other high risk study populations and highly defined populations such as prisoners, HIV/STD clinic attendees indicating recently acquired sexually transmitted disease, sex workers, multi-transfused patients, drug addicted individuals and injecting drug users, liver transplant recipients, refugees, homeless people.

(c) Reports of acute diseases surveillance (reporting incident cases, acute cases or rates per 100,000 population).

(d) Reports and data that are incomplete (e.g. number tested or not provided or below 20).

(e) Citations reporting HBsAg prevalence in the US other than NHANES prevalence.
Prevalence of hepatitis B infection, children 5-9 years, 2005

Map 1. Map for children.

Prevalence of hepatitis B infection, adults 19-49 years, 2005

Fig. 2–5. HBsAg seroprevalence by region and age-group for males and females separately, 1990 and 2005.
2005 compared to 3.9% in males. We estimated 240 million people chronically infected worldwide in 2005 (Table 1).

### Table 1
Overview: Global HBsAg and people chronically infected.

<table>
<thead>
<tr>
<th>Year</th>
<th>Males Persons HBsAg positive</th>
<th>Prevalence</th>
<th>Females Persons HBsAg positive</th>
<th>Prevalence</th>
<th>Both Persons HBsAg positive</th>
<th>Prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>118 million</td>
<td>4.4</td>
<td>105 million</td>
<td>4.0</td>
<td>223 million</td>
<td>4.2</td>
</tr>
<tr>
<td>2005</td>
<td>127 million</td>
<td>3.9</td>
<td>113 million</td>
<td>3.5</td>
<td>240 million</td>
<td>3.7</td>
</tr>
</tbody>
</table>

3.1. HBsAg prevalence by GBD world region (Tables 2–5: Web Annex 3)

HBV prevalence was most common in sub-Saharan regions of Africa. Western sub-Saharan African countries had some of the highest age-specific HBsAg prevalence in the world reaching up to 12% among children and adolescents in the age-groups up to 19 years in 1990. Although there was a decrease in 2005, the region continued to have high HBV endemicity, which is more pronounced among males.

An increase in chronic HBV infection among younger age-groups (0–14 years) occurred in Southern sub-Saharan Africa in 2005 compared to 1990 that resulted in age-specific prevalence of 8–9% among young females. Also, Eastern sub-Saharan African countries faced an increase in the youngest ages and almost no change in other age-groups. In 2005, prevalence peaked at approximately 7% in 0–4 years aged boys and girls and declined with age in this region.

A decrease in prevalence was evident in Central sub-Saharan Africa which transitioned from high endemicity among younger individuals (age-groups up to 34 years) in 1990 into intermediate endemicity across all ages in 2005.

North Africa and the Middle Eastern region showed lower intermediate HBsAg endemicity across all age-groups in 2005. Prevalence decreased from 1990 to 2005, particularly among males up to 34 years.

The prevalence in high-income countries of North America (Canada and the United States) was low and declined among both sexes and across all ages between 1990 and 2005. Males had higher HBsAg positivity than females in both periods, peaking in the male 0–4 years age-group at 2.71% and 2.14% in 1990 and 2005, respectively. The oldest ages (65+ years) showed the lowest prevalence of approximately 1% in 2005.

Both Tropical Latin America and Central Latin America demonstrated a strong decrease in HBsAg prevalence between 1990 and 2005. Tropical Latin America changed from an intermediate into a low endemicity region. Where 0–9 year aged boys had a higher intermediate endemicity of over 5% in 1990, HBsAg prevalence was only 1.6% in 2005. Similarly, in Central Latin America prevalence has halved in this period and most adult age-groups shifted to a low endemicity level in 2005. Other Latin American regions such as Andean Latin America and Southern Latin America showed a decreasing prevalence by age but relatively constant intermediate endemicity levels. A slight decline in prevalence from 1990 to 2005 among Andean Latin Americans was paralleled by an increase in HBsAg prevalence in Southern Latin America.

HBV chronic infection rates in Caribbean children and adolescents aged 0–19 years ranged from 4.3% to 5.4% and was fairly constant over time. HBsAg prevalence decreased with age.

In the island nations of the Pacific and Indian Oceans (Oceania), HBsAg was highly endemic among men and women in 1990, peaking at approximately 10% in men aged 10–34. The decrease in prevalence up to 2005 led to a shift into a higher intermediate endemicity level among the age-groups up to age 54 and into a lower intermediate endemicity level in older adults (55+).

Among the European regions, seroprevalence of HBsAg showed consistently low prevalence in Western Europe. This was particularly true for females who had prevalence below 2% throughout the time periods. Nevertheless, between 1990 and 2005, an increase in both sexes was observed that led to a change from initially low endemicity in young males to a low-intermediate endemicity level in 2005; this was accompanied by a decrease in prevalence in the older individuals (65+ years). Central and Eastern European children had a higher intermediate HBsAg endemicity in 1990, which decreased in 2005. Older Central European females in particular demonstrated a strong HBsAg prevalence decrease up to 2005. Prevalence in infant and young girls declined from 6% in 1990 to 3% in 2005 in this region. In contrast, Eastern European countries did not experience as strong a reduction in HBsAg prevalence in the youngest age-groups. In both Central and Eastern Europe the age group 0–9 years remains the most affected by HBsAg infection.

Among all Asian regions, East Asia had the highest prevalence of HBV infection and there was not much of a change between 1990 and 2005 apart from a slight decrease in children and an increase in all age-groups above 25 years in 2005 as compared to 1990. In 1990, prevalence decreased in both sexes with age, but increased with age in 2005 and showed the highest prevalence of over 8% among males aged over 35 years. Generally, endemicity remains at high or close to a high endemicity in this region, which is particularly true for males. Intermediate HBsAg endemicity was also estimated for Central Asia, which includes the Caucasus and central Asian countries. A small decrease was observed between 1990 and 2005 but Central Asian children and younger adults had an HBsAg prevalence of around 5% in 2005. In South Asia, approximately 3% of the population up to age 45 was HBsAg positive with a decrease in older individuals who demonstrated a low prevalence in 2005.

Unlike other Asian regions, South East Asia experienced a strong reduction in HBsAg prevalence between 1990 and 2005, particularly in the young age groups of 0–14 years that had prevalence levels of 1.2–1.4% in 2005. In contrast, South East Asian adults appeared to continuously have higher-intermediate HBsAg prevalence of 5% to over 6% in 2005. The pattern of an increasing prevalence with age was very pronounced in 1990 and rather exceptional compared to the other regions.

Some reduction in HBsAg prevalence between 1990 and 2005 occurred also in high-income Asian Pacific countries including Japan, the Republic of Korea, and Singapore. The middle ages (25–54) were the most affected age groups but overall, endemicity remained at an lower intermediate level of approximately 4% in 2005. Interestingly, in 1990 the oldest age-groups of over 75 years had the highest prevalence among all age-groups. Australian countries experienced a reduction in HBsAg prevalence and were categorized as a lower intermediate endemicity region in 2005. Males of all age-groups had prevalence in the range of 4% in 2005 whereas 2–3% of females up to age 55 were affected by chronic HBV infection with a sharp increase in the oldest age-groups.

4. Discussion

We found a very large burden of HBsAg infection in all sub-Saharan African regions, East Asia and, to a lesser extent, in Oceania
and Andean Latin America. Most other regions with high and middle income showed a mix of lower and higher intermediate HBV endemicity. Only a few regions demonstrated prevalence below 2% throughout most age-groups. Among these were Tropical Latin America, Central Latin America, North America and Western Europe. There was an overall decrease in HBsAg prevalence from 1990 to 2005 in younger age-groups, which may be closely related to widespread hepatitis B immunization, particularly in low income regions. Significant decreases in HBsAg prevalence due to immunization were reported from African countries such as the Gambia [15,16] and Senegal [17]. However, the infection remains extremely prevalent in sub-Saharan Africa and the attributable HBV-related disease burden can be expected to remain high. This is also reflected in a high mortality from primary liver cancer, one sequela of chronic HBV infection and the most frequent cause of cancer deaths among men in this region [18].

The observed decreases in HBsAg prevalence in North America and Europe were temporally associated with increased hepatitis B vaccine coverage rates [19], improved screening of blood products and increased availability of safe injection materials [20,21]. The generally low HBsAg endemicity levels in these countries are paralleled by a steady decline in reported cases of acute hepatitis B [22,23]. On the other hand, the large number of individuals infected with HBV influences the number of liver cirrhosis and cancer cases in many world regions including high income countries [24] and highlights a need for screening and surveillance programs to identify chronically infected individuals and thereby prevent further transmission as well as to provide opportunities for secondary and tertiary prevention [3].

Asian countries, for example those in the GBD South East and East Asian regions, have also experienced dramatic increases in coverage of routine infant hepatitis B vaccine that were accompanied by a reduction in HBsAg prevalence. The impact of this decrease in HBsAg prevalence was also measurable in a substantial reduction in the HBV-related disease burden in countries that were highly endemic in the past e.g. Taiwan [25,26] and China [27,28].

Strengths of this study include the extensive systematic literature review and the use of an empirical Bayesian hierarchical model to estimate region-specific HBsAg prevalence and endemicity and to subsequently calculate the absolute number of people being chronically infected with the HBV.

This study has a number of limitations. Observed HBsAg prevalence data are lacking in some regions and the quality of studies reporting these data is often low. Middle- and low income regions, e.g. Oceania, Central Asia, and Andean Latin America had a limited evidence base or studies were concentrated on one country as is the case with India as part of the South Asian region or Thailand located in the South East Asian region. Accordingly, simulations of prevalence may lead to potential underestimation of the true regional profile, particularly if studies were more likely to be conducted in countries with higher economic standards and better research infrastructure. To address issues of representativeness, grey and non-English literature should be considered in future studies and there is a need for generating more high quality data from low resource settings. Most high quality studies were conducted in high income countries, for example the Western European region. Nationally representative, population-based studies reporting HBsAg prevalence were only available from the US.

Another limitation is related to factors that were not considered in our analysis such as genotype information. It is known that viral genotypes vary between and within countries, depending on the populations at risk and their geographical origin. Very few HBV prevalence studies report on genotype and the genotype-specific HBV distribution by country has not been sufficiently studied. As a result, we did not adjust for this information, which might be crucial given the fact that some genotypes are associated with more severe disease or clinical response to treatment [29–31]. Similarly, the laboratory method used to detect HBsAg was only reported in a few studies and we did not consider this factor in the analysis, which could impact the comparability of HBsAg prevalence across studies.

Since the overall objective of this study is to provide a regional picture of HBV prevalence, the results do not capture the potential heterogeneity that exists between sub-populations within a country. It should be noted that some low endemicity areas in Western Europe and North America face great intra-country variation with higher prevalence and higher hepatitis-related mortality among migrants [32–35] and additional country-specific data would be crucial for comprehensively guiding national hepatitis B prevention and control programs and targeting most vulnerable population groups.

Prevalence data obtained from systematic reviews and modeling should be interpreted conservatively. Descriptive epidemiological research conducted in high income areas may generally focus more on marginalized and higher risk populations whereas studies from low income regions may focus on urban and higher educated populations that experience lower infection rates than those living in poorer areas. As a result, the prevalence of chronic HBV infection reported in this study could be underestimated for high income regions and overestimated for low income regions. This would, however, increase the estimated differences between these regions and support our findings.

There is a need for systematically collected and population-based HBV infection data. Data on other markers of HBV such as HBsAg and anti-HBc are also needed to describe current and future HBV-related disease burden.

Acknowledgements

We greatly acknowledge the contribution and efforts of Abraha Flaxman for providing invaluable guidance on using and adapting Dismod III, and Seth Flaxman for adapting Dismod III for our purposes. We greatly acknowledge the efforts of Don Ward († Deceased September 7, 2010) and his team who abstracted the studies. We thank Erica Din, Craig Lammert and Melissa Creary for their work with searching, abstracting and organizing data. We acknowledge the assistance of Claire Preaud for geographical maps.

Contributors: JJO wrote the manuscript, modeled, analyzed, and interpreted the data. GAS provided guidance on data modeling and contributed to writing. JG designed and conducted the literature search and reviewed articles. STW initiated the study, supervised all components of the study and contributed to writing the manuscript.

Conflict of interest statement: None declared.

Funding: We acknowledge the financial support of the Centers for Disease Control and Prevention’s Division of Viral Hepatitis, the World Health Organization and the Bill and Melinda Gates Foundation.

Appendix A. Supplementary data


References


