Online technical appendix

Global hepatitis report 2026

WHO approach for producing global and regional viral hepatitis estimates: data sources and methods

World Health Organization

Online technical appendix

WHO approach for producing global and regional viral hepatitis estimates: data sources and methods.

1. Introduction

To support the preparation of the Global hepatitis report 2026, the CDA Foundation provided the World Health Organization (WHO) with hepatitis datasets from different levels – national, regional and global – generated through the Polaris Observatory. The modelling analyses from the Polaris Observatory produce standardized, population-level estimates of hepatitis B and C prevalence, incidence, mortality and cascade-of-care outcomes at national, regional and global levels. These analyses cover years 2015 to 2024, and include forward projections to 2030 when available.

For this report, WHO received a consolidated dataset containing global and regional estimates for the period **2015–2024**. The modelled outputs include annual estimates of prevalence, incidence, mortality and key cascade-of-care indicators (including numbers diagnosed and treated) for both hepatitis B virus (HBV) and hepatitis C virus (HCV). Definitions of all indicators and population groups are provided in **Table A1**.

This annex provides technical notes on the modelling methods used to develop these estimates. It is intended to support transparency, reproducibility and alignment with WHO's standards for the reporting of global health estimates.

All country-level datasets were shared with national focal points as part of the global reporting process and were also incorporated into the WHO country consultation mechanism. Further details on the consultation and validation steps with Member States are provided in **Annex 2**.

2. Overall approach of the modelling methods

Methodological details of the HBV and HCV models, including key parameters and indicator definitions, are provided in **Table A1** and in **Sections 4–5**, and in the corresponding peer-reviewed references *(1-5)*. The CDA Foundation's modelling approach is grounded in a systematic review of biomarker surveys, which provides the empirical basis for estimating HBV and HCV prevalence. Using these data, country-specific models were constructed to generate annual estimates of incidence, prevalence and mortality to the end of 2024. Cascade-of-care estimates were derived from country-level model outputs, complemented by pharmaceutical sales data, expert input from national stakeholders and extrapolation procedures when direct data were unavailable. To ensure stability of long-term disease trajectories, historical simulations were run from the mid-20th century (1950) to the end of 2050. Annual outputs for the period 2015–2030 were extracted for global reporting.

The model structures were as follows:

> **HBV model structure (PRoGReSs model)**: the HBV model uses a dynamic mathematical framework that estimates annual HBV prevalence and incidence by stage of liver disease, age and sex using disease progression and mortality (all-cause and liver-related) rates. The model integrates demographic profiles, vaccination coverage (including timely birth dose and three-dose infant

series), prevention measures such as hepatitis B immune globulin (HBIG) and maternal antiviral prophylaxis, HBV diagnosis and treatment schedules, and established HBV epidemiological parameters. Perinatal, early childhood and horizontal transmission are modelled using age-specific hepatitis B surface antigen (HBsAg) prevalence together with available prevention-coverage data. This allows the model to reflect both historical and current programme performance. Full technical documentation of the PRoGReSs HBV model have been previously published *(6)* and are summarized in section 5 .

HCV model structure: the HCV model employs a Markov/semi-dynamic natural history framework that follows people with viraemic infection across stages of liver disease, by age and sex, over time using disease progression and mortality (all-cause and liver-related) rates. It incorporates inputs from demographic profiles, HCV diagnosis and treatment initiation schedules, and subsequent cure (sustained virological response, SVR). Full technical documentation of the HCV model have been previously published *(1)*and are summarized in section 5.

3. Data inputs

The country-level estimates used in this report were generated through a two-step process: **(1) systematic review** of epidemiological data, followed by **(2) model-based estimation** using country-specific HBV and HCV models. Inputs were compiled through comprehensive searches of published and unpublished sources, supplemented by expert validation and contextual information obtained through a structured Delphi process. **Table A2** provides details on the Delphi process.

Systematic literature review – HBV

A systematic review was conducted to identify all available hepatitis B biomarker prevalence studies. Searches were performed in PubMed and Embase for research published between 1 January 1960 and 1 March 2016, without language restrictions, using the designated terms. The review was updated to include studies published between 1 March 2016 and 1 December 2024 *(3)*. Titles and abstracts were screened for relevance, and only studies reporting HBsAg prevalence were included. In addition to peer-reviewed evidence, the review incorporated grey literature, ministry of health (MoH) reports, conference abstracts, local journals and unpublished data shared by national experts. Studies were assessed for representativeness and methodological quality before being incorporated into country-specific HBV models *(3, 4)*.

Systematic literature review – HCV

A parallel systematic review was undertaken for HCV. As with HBV, PubMed and Embase were searched without language restrictions for research published in the period 1 January 1960 to 1 March 2016 using combinations of country names with terms for HCV prevalence, viraemia, genotype distribution and epidemiology. The search was updated to include new evidence published through 1 December 2024, ensuring consistency with the HBV search window. To support the HCV model inputs, the review identified studies that reported anti-HCV antibody prevalence, HCV RNA (viraemic) prevalence, age- or sex-specific distributions, genotype profiles and fibrosis-stage distribution where available. As with HBV, relevant grey literature, programme reports, blood-donor surveillance and expert-shared unpublished datasets were included when judged to be robust and representative.

3.1 Inclusion and exclusion criteria (HBV and HCV model inputs)

For both viruses, studies were included only if they met preset quality criteria regarding population representativeness and methodological validity.

Included: general-population surveys; large, representative subpopulations (e.g. antenatal clinic attendees and national school-based surveys); national or multisite biomarker surveys; and studies that clearly reported sample size, testing method, year(s) of data collection and population characteristics.

Excluded or used cautiously: highly specific clinical subgroups (e.g. haemodialysis patients, liver clinics and people who inject drugs – unless used for specific model components such as fibrosis distribution), studies without laboratory-confirmed infection markers, and studies lacking sampling methods or with unclear representativeness.

Prevalence data were country-representative where possible. Non-representative populations (e.g. selected clinical groups) were generally excluded or used cautiously. Where no suitable country-specific data were available, the modelling framework applied population-weighted regional averages to fill the gaps.

Variables extracted from included studies

For HBV, the following variables were systematically extracted:

- HBsAg prevalence;
- age and sex distribution of tested populations (when available);
- study year(s) and testing methods;
- population type (e.g. general population, pregnant women and students);
- study setting (urban/rural) and geographic scope (single site, multicity, regional and national)
- study design category (surveillance, cross-sectional survey, meta-analysis, modelling study and review); and
- sample size and sampling approach.

These variables informed inputs for HBV transmission dynamics; vertical, early childhood and horizontal infection probabilities; and age-stratified prevalence curves.

For HCV, extracted variables paralleled those for HBV but additionally included HCV-specific attributes:

- anti-HCV antibody prevalence and HCV RNA (viraemic) prevalence;
- genotype distribution by age, sex or risk group;
- stage-of-disease indicators (fibrosis staging or aspartate aminotransferase [AST] to platelet ratio index/fibrosis-4 [APRI/FIB-4] proxies, when available);
- time period of data collection and diagnostic assays used;
- population characteristics (general population, blood donors, antenatal women and community surveys);
- study representativeness, geographic coverage and sampling methodology; and
- sample size and study design (survey, surveillance data analysis and meta-analysis).

These variables informed model calibration, estimation of viraemic prevalence and construction of fibrosis-stage distributions.

Scoring process for the extracted data identified in literature search: extracted data from literature were scored using a multi objective decision-analysis approach (see **Box A1**), resulting in a score of 1–3 for each study. For published studies, the overall score was based on the weighting of the scores for generalizability, sample size and year of the study. The highest scoring study was chosen to provide the representative estimate for each country, with the exception of countries in which lower scoring studies were recommended by local experts. Although blood-donor studies were excluded from use as base estimates, they were used to provide estimates of the lower bound of the uncertainty intervals (UIs) (Section 3.2).

Expert consultation and Delphi validation

For both HBV and HCV, systematic review outputs were supplemented by structured Delphi processes conducted with national and regional experts. Experts contributed to:

- identifying additional unpublished or non-indexed data sources;
- validating study relevance and representativeness;
- clarifying discrepancies between datasets;
- endorsing final country-specific model inputs; and
- determining appropriate analogues for extrapolation when data gaps existed.

Full details of the Delphi process are provided in **Table A2**.

3.2 Development of uncertainty intervals

UIs for the HBV and HCV estimates were developed through a structured process grounded in the quality, variability and availability of prevalence data. For each country, all identified biomarker prevalence studies were scored using predefined quality criteria. The highest scoring study was selected as the base value for the model. In settings where country experts recommended an alternative study – typically owing to stronger contextual relevance or more recent field experience – those expert-validated studies were used instead. Although blood-donor studies were excluded from serving as base prevalence estimates, they were retained as informative lower bound inputs in the uncertainty analysis because of their generally lower prevalence and large sample sizes.

For the most part, UIs were generated around prevalence rather than for every input parameter. For HBV and HCV, the high and low values of the UIs were derived by comparing the base prevalence study with all other eligible prevalence sources in the country. Where multiple studies existed, the highest and lowest plausible estimates were used to define the uncertainty bounds. In countries or territories with marked subnational heterogeneity, the expert panel recommended the use of population-weighted averages across available regions to generate the base value. In such cases, the lowest prevalence region was retained to anchor the lower bound, and the highest prevalence region served as the upper bound.

For HCV, viraemic prevalence uncertainty was derived by pairing anti-HCV prevalence UIs with the range of viraemia proportions observed in the literature and in regionally comparable countries. Similarly, for HBV, uncertainty around HBsAg prevalence incorporated variation in study methodologies, sample selection and testing approaches. Incidence, mortality and cascade-of-care indicators (diagnosis and treatment for HBV, diagnosis only for HCV) are *model-derived* through deterministic transitions; hence,

the model does not compute a unique UI for each of these indicators. Instead, the model propagates the uncertainty around prevalence – the major empirical driver of the disease burden – resulting in high and low trajectories for downstream indicators. These prevalence-based UIs are then reflected in estimated ranges for total infections, chronic disease stages and liver-related mortality. For HCV, separate UIs for treatment data were available *(7)*.

Overall, the UIs represent the **plausible range of prevalence and treatment data consistent with the available evidence**, including both published and country-provided data. They reflect variation in study quality, sampling frames and geographic heterogeneity, as well as expert judgement where empirical data are sparse. This approach aligns with previous Polaris Observatory modelling work, and ensures that the UIs used in the *Global hepatitis report 2026* are transparent, evidence-based and consistent for both HBV and HCV.

3.3 Potential biases in model inputs

Despite the use of comprehensive systematic reviews, grey-literature searches and structured expert validation, several inherent limitations and potential sources of bias affect both the HBV and HCV model inputs, as outlined below.

General limitations

- **Heterogeneity and gaps in general population data**: many countries lack recent, nationally representative HBV or HCV biomarker surveys. Available studies often vary in methodology, population coverage and diagnostic assays.
- **Selection bias in available studies**: where only subpopulation studies exist (e.g. antenatal clinics, blood donors and clinical cohorts), these may not accurately reflect national prevalence in the general population.
- **Underrepresentation of small countries and data-poor settings**: in such contexts, model inputs rely heavily on regional averages or expert elicitation, increasing uncertainty.
- **Reliance on expert input where empirical data are scarce**: expert consensus helps to fill critical gaps but may introduce subjective bias, particularly for historical prevalence, treatment uptake and age-stratified distributions.
- **Limitations of literature-based and secondary data**: even comprehensive reviews cannot fully capture variations in transmission dynamics, service access and population movements.

Because of these limitations, WHO performs an additional **regional and national validation step through the WHO country consultation process**. This step is essential for achieving consensus on model inputs, ensuring alignment with national programme data, and reconciling discrepancies between modelled estimates and country-derived figures.

Biases and limitations specific to HBV: for HBV, several additional factors may influence the accuracy of model inputs, as outlined below:

- **Incomplete measurement of high-risk or underserved populations**: HBV prevalence may be underestimated when data do not sufficiently represent populations at higher risk (e.g. migrants, Indigenous peoples, rural communities or regions with historically low infant vaccination coverage).
- **Uneven access to prophylaxis and vaccination**: models do not explicitly quantify rural–urban disparities in birth dose availability, facility deliveries or HBIG access, potentially masking within-country inequities.
- **Migration not fully accounted for**: immigration from high-prevalence regions may increase HBV burden in countries with otherwise low endemicity; this impact is not explicitly modelled; and
- **paediatric estimates in data-poor settings**: estimates for children aged 5 years or below rely heavily on modelled estimates when no serosurveys or programme data exist, which may underestimate or overestimate true prevalence.

Biases and limitations specific to HCV: similarly, there are several limitations affecting HCV estimates:

- **underrepresentation of key populations in national data**: people who inject drugs, incarcerated populations and other high-risk groups often lack nationally representative data; when these populations carry a disproportionate burden of HCV infection, national models may underestimate overall viraemic prevalence.
- **Scarcity of recent viraemic (HCV RNA) data**: many countries have anti-HCV antibody studies but lack RNA-based estimates that measure active infection; hence, there is a need for model-based conversion using historical viraemia ratios.
- **Limited data on fibrosis stage and treatment cascades**: estimates of chronic liver disease burden and treatment eligibility depend on sparse or outdated fibrosis-stage data, especially where routine clinical staging is limited.
- **Impacts of service disruptions not systematically quantified**: variations in harm-reduction coverage, testing policies and treatment availability – particularly following coronavirus disease (COVID-19) disruptions – may not be fully reflected in the input datasets.

Geographic and demographic gaps: for both HBV and HCV, national-level estimates may obscure:

- substantial **subnational heterogeneity**;
- differences between **rural and urban settings**;
- regional disparities in **health service access**, including vaccination, prevention of mother-to-child transmission (PMTCT), testing and treatment; and
- differences in burden among **mobile or marginalized populations**.

These gaps highlight the continuing need for **nationally representative biomarker surveys**, strengthened routine surveillance and improved reporting of high-risk populations.

Key assumptions and transformations

The HBV and HCV models apply a set of shared analytical assumptions to harmonize inputs and maintain comparability across countries, using:

- **prevalence calibration**: both models prioritize representative national surveys; when unavailable, adjusted grey-literature sources (including surveys of blood donors or antenatal clinics) are triangulated cautiously to approximate population prevalence;
- **cascade-of-care metrics**: estimates of diagnosis and treatment incorporate country-reported programme data, pharmaceutical market data and expert review; for HCV, **scale-up of direct-acting antivirals (DAAs)** is explicitly modelled, allowing cured cases (SVR) to be removed from the viraemic pool; and
- **regional borrowing**: when a country lacks sufficient primary inputs, **population-weighted United Nations (UN) subregional averages** are applied to derive prevalence, viraemia ratios, treatment initiation rates and intervention coverage.

These assumptions ensure comparability across countries while acknowledging the uncertainties arising from incomplete or heterogeneous data sources.

Table A1. Summary of model parameters and data sources

Model input	Definition	Source
Population and demographic parameters		
Population, annual, by sex and age	Mid-year population distribution used for age-specific prevalence, transmission and mortality estimates	*(8)*
Background mortality, annual, by sex and age	All-cause mortality	*(8)*
Births or fertility, annual, by maternal age	Number of births or number of births per woman (for modelling perinatal transmission)	*(8)*
Sex ratio at birth, annual	Ratio of male to female births	*(8)*
HBV-specific model inputs (PRoGReSs HBV model)		
(i) Diagnosis and treatment inputs		
Newly diagnosed, annual	Number of new HBV diagnoses	*(3, 5)*
On-treatment population, annual	Number of HBV-infected individuals receiving antiviral therapy	*(3, 5)*
Treatment discontinuation rate, annual	Proportion of those receiving treatment discontinuing antiviral therapy	*(5)*
On-treatment virological suppression, annual	HBV DNA suppression rate for individuals adherent to treatment	*(9)*
(ii) Prevention inputs (MTCT and childhood)		
Maternal antiviral prophylaxis coverage, annual	Proportion of pregnant women who are HBsAg-positive or have high viral load receiving antiviral prophylaxis	*(5)*
HBIG use, annual	Proportion of infants receiving HBIG, where indicated	*(5)*
Timely birth-dose vaccination coverage, annual	Proportion of neonates receiving a birth dose within 24 hours	*(10)*
3-dose infant vaccination coverage, annual	Proportion of infants who have received complete, or partial, HBV vaccine series	*(10)*
Catch-up vaccination coverage, annual, by 5-year age group, annual	Proportion of non-infants who have received complete HBV vaccine series	*(10)*
(iii) HBV natural history and transmission parameters		
HBsAg prevalence target, by age or sex, in a given year	Calibrated age- or sex-specific chronic infection prevalence	*(3, 5)*

Model input	Definition	Source
HBeAg-positive among HBsAg-positive women of childbearing age, in a given year	Proportion of HBsAg-positive women of childbearing age who are HBeAg-positive	Country-specific data
HBeAg-negative with high viral load, in a given year	Proportion of HBeAg-negative population with HBV DNA ≥20 000 IU/mL	*(11, 12)*
HBeAg-positive with high viral load, in a given year	Proportion of HBeAg-positive population with HBV DNA ≥20 000 IU/mL	*(11-13)*
Perinatal transmission rates of HBV by vaccination status of infants, and maternal serologic and treatment status	Transmission probabilities stratified by maternal serologic status, viral load, and vaccination and HBIG coverage of infants	(**Table A3**)
Risk for chronic infection, for perinatally acquired HBV infection, $c_{\text{Perinatal}}$; and for horizontally acquired HBV infection at age a, c_a	Proportion of incident cases of HBV infection who remain chronically infected with HBV $c_{\text{Perinatal}} = 0.885$ $c_a = \begin{cases} c_1 & \text{if } a = 0 \\ 1 - 0.7145a^{0.0814} & \text{if } 1 \leq a < 35 \\ c_{34} & \text{if } a \geq 35 \end{cases}$	*(14)*
Annual progression rates, start health state → end health state	Proportion of HBV-infected population progressing from start health state to end health state annually	*(5)*
HCV-specific model inputs (HCV Markov/semi-dynamic model)		
(i) Epidemiological inputs (prevalence and viraemia)		
Anti-HCV prevalence, in a given year	HCV antibody positivity rate in the general population	*(2, 5)*
Viraemic prevalence, in a given year	Proportion of anti-HCV(+) individuals who are HCV RNA+	*(2, 5)*
Viraemia conversion factor	Empirical ratio used where RNA data are unavailable	*(2, 5)*
(ii) Diagnosis and treatment cascade		
Treated, annual	Number receiving antiviral therapy	National reports, published studies, drug-sales adjustments
Newly diagnosed, annual	New HCV diagnoses	National registries OR population-weighted regional averages
Total previously diagnosed, in a given year	All viraemic individuals ever diagnosed and still living	National registries OR Polaris extrapolation

Model input	Definition	Source
SVR (cure) rate, annual	Antiviral treatment cure probability	*(2, 5)*
(i) HCV natural history and progression		
Annual progression rates, start health state → end health state	Proportion of HCV-infected population progressing from start health state to end health state annually	Multicountry expert review (Polaris)
Shared HBV/HCV sequelae and advanced liver disease indicators		
Liver transplantations, annual	Liver transplantations attributed to HBV/HCV infection	IRODaT, national registries adjusted for viral attribution *(15)*
HCC, annual	HCC incidence attributed to HBV or HCV infection	GLOBOCAN, national cancer registries
Cirrhosis prevalence and mortality	Burden of decompensated and compensated cirrhosis	Polaris HBV/HCV models
Liver-related mortality	Deaths from cirrhosis/HCC attributable to HBV or HCV infection	Polaris estimates, GLOBOCAN
Model calibration and assumptions (applies to both HBV and HCV models)		
Incidence multipliers	Back-calculated incidence estimators used for prevalence–incidence consistency	Polaris HBV/HCV calibration (**Section 4**)
Regional borrowing	Use of population-weighted regional averages where data gaps exist	Polaris HBV/HCV calibration (**Section 4**)
Expert-validated inputs	National expert Delphi process to confirm or adjust model parameters	Polaris Delphi methodology (**Table A2**)
Historical back-casting	Long-term simulation (1950–2050 for HCV; multidecade for HBV) to stabilize prevalence trajectories	Polaris HBV/HCV calibration (**Section 4**)

HBeAg: hepatitis B e antigen; HBIG: hepatitis B immune globulin; HBsAg: hepatitis B surface antigen; DNA: deoxyribonucleic acid; HBV: hepatitis B virus; HCC: hepatocellular carcinoma; HCV: hepatitis C virus; IRODaT: International Registry of Organ Donation and Transplantation; MTCT: mother-to-child transmission; RNA: ribonucleic acid; SVR: sustained virological response.

Table A2. Delphi process

		Activities
Phase 1 – data gathering	1a	**Identify country experts who are willing to collaborate** • Experts were identified through HBV-related scientific contributions, or through referrals and recommendations from leading researchers. Panels consisted of hepatologists, gastroenterologists, virologists, infectious disease specialists, epidemiologists, health economists, health scientists and MoH representatives.
	1b	**Undertake a literature search** • Review the internal database for previously identified sources. • Review online sources (e.g. MoH and WHO) to capture non-indexed sources. • Run a literature search from 2020 forward to identify recent publications. • Summarize input data available through the literature. • Gather empirical data for new HCC cases, LTs, percentage of HCC and LT due to HBV, annual number of people newly diagnosed and annual number of people treated. • Build a draft model based on published data or extrapolate inputs from countries with data when data are missing (as a placeholder). • Schedule a meeting with experts.
Phase 2 – country meetings and modelling	2a	**Hold an initial expert meeting (Meeting 1, 2–3 hours)** • Provide background information on the project, model and methodology. • Review data identified in Phase 1b and highlight gaps in data. • Request data from local non-indexed journals, unpublished data and any other available data (e.g. hospital-level data) that can be used to fill the gaps. • Gain agreement on countries that can be used for extrapolation when no local data are available.
	2b	**Follow up with experts after Meeting 1** • Send minutes of Meeting 1 and a list of remaining action items to the experts identified in Phase 1. • Follow up with the experts to collect missing data and obtain copies of publications in local journals, unpublished data, relevant PhD theses, government reports and raw hospital or registry-level data. • Analyse raw data and send the results to the experts for approval.
	2c	**Model the disease burden** • Populate the disease burden model with inputs and calibrate model to empirical data. • Develop 2–3 scenarios to prepare for Meeting 2, including a WHO target scenario (elimination by 2030). • Schedule the second meeting. • Develop a slide deck summarizing all inputs and associated data sources. • Perform a final check of the model and slide deck and approve internally.

		Activities
	2d	**Hold a second expert meeting (Meeting 2, 2–3 hours)** • Review all inputs as well as data provided by the experts since Meeting 1 and the results of analyses of any raw data provided. • Gain agreement on all inputs to be used in the model. • Update the model using any updated inputs. • Run scenarios requested by the experts (e.g. slow increase in the number of treated patients, disease control and WHO target) and review the results and insights. • Agree on final strategies that could be considered as part of a national strategy.
Phase 3 – follow-up analyses	3a	**Undertake follow-up analyses** • Update the model as necessary and send the results to the experts. • Provide support to address follow-up questions. • Establish inputs and outputs as approved. • Run additional scenarios to support the development of a national strategy (e.g. economic impact, birth cohort screening and sources of transmission). • Report the results to the Polaris Observatory. • Update the analysis as new information becomes available (e.g. new national studies and updated treatment data).

HBV: hepatitis B virus; HCC: hepatocellular carcinoma; LT: liver transplant; MoH: ministry of health; WHO: World Health Organization.

Box A1. Scoring process

HBV prevalence studies were scored on a scale of 0–10, following the approach described previously *(3, 4)*. This system was based on three metrics, which accounted for 60% (generalizability), 20% (sample size) and 20% (year of analysis) of the overall score, respectively:

$$\text{Overall score} = 60\% \times \text{Generalizability score} + 20\% \times \text{Sample size score} + 20\% \times \text{Year of analysis score}$$

Sample size score: The log of sample size was scaled from 0 to 10, whereby all studies with a sample size greater than 10 000 received a score of 10.

Year of analysis score: The study year was assessed so that analyses conducted after 2019 received a score of 10; 2015–2019, a score of 8; 2011–2014, a score of 6; 2004–2010, a score of 4; 2000–2003, a score of 2; and analyses conducted before 2000 a score of zero. All studies that did not report the year of study were assumed to have been conducted 2 years before the year of publication.

For simplicity, the 0–10 scores were converted to a data quality scale of 1–3, where an overall score of 0.0 to <4.0 received a data quality score of 1, an overall score of 4.0 to <8.0 received a quality score of 2 and an overall score of 8.0 to <10.0 received a quality score of 3. Modelling studies were automatically given a data quality score of 2. Studies without a formal assessment, but deemed to be of quality for inclusion, were given a score of 1.

Generalizability score: The grid below indicates the criteria used to score articles on their ability to be generalized to the total population.

Expert consensus was assigned a default score of 1, unless supportive data were available. Expert consensus estimates based on supporting data were scored as follows: 2 = expert input based on published or unpublished data; 3 = expert consensus based on well-conducted studies ahead of print or large national databases.

Geographical scope ↓	**Scale, 0–10**				
					10[†]
National	0	3	4 Meta-analysis: 4	6	9 Model: 6 Meta-analysis: 5
Large region **Multi-region** **Multicity** **Large city**	0	1	2–3[a]	4–5[a]	6–8[a]
Small region/town **village** **tribe** **hospital**	0	0	1	1	2
Population →	**High risk, any sampling method** **- People who inject drugs** **- People with HIV** **- Surgical patients** **- Blood transfusion recipients**	**Healthy adults, self-selected** **- Blood donors**	**Healthy adults, self-selected** **- Health checkup patients** **- Screening**	**Healthy adults, randomly selected** **- Health care workers** **- Pregnant women** **- Soldiers**	**General population, randomly selected**

[a] Reserved for a nationally representative sample with a stratified, multistage and random sampling design that documents the study design and demographics of subjects thoroughly; variability subject to author's discretion based on quality of study design and the geographic scope of the respective country.

4. Method of computation

4.1 HBV modelling methods and inputs

The PRoGReSs model was named after the modellers who developed it: Ken **P**asini, Homie **R**azavi, Ivane **G**amkrelidze and Devin **R**azavi-**S**hearer. It is a compartmental, deterministic, dynamic Markov disease progression model developed in Microsoft (MS) Excel® and MS Visual Basic (Microsoft Corporation, Redmond, WA, United States of America [USA]) to quantify the annual HBV-infected population by disease stage, sex and age in a country. MS Excel was selected because of its transparency, flexibility and widespread availability. Disease progression rates are available in the most recent publication *(3)*.

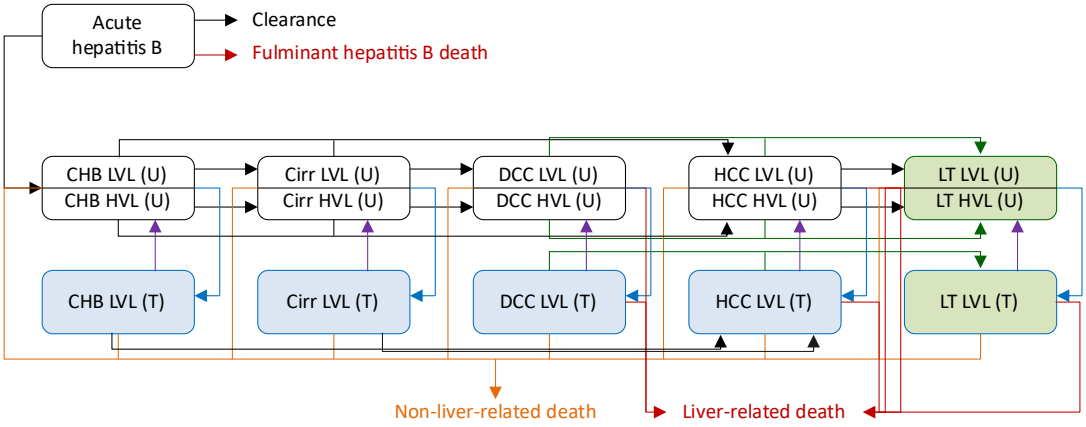
The disease stages considered in the PRoGReSs model were chronic hepatitis B, compensated cirrhosis, decompensated cirrhosis, hepatocellular carcinoma and liver transplantation. Populations with decompensated cirrhosis and hepatocellular carcinoma were considered to be eligible for a liver transplant.

The HBV-infected population in each disease stage was further divided into high viral load (HBsAg-positive with HBV DNA of ≥20 000 IU/mL), low viral load (HBsAg-positive with HBV DNA of <20 000 IU/mL) and treatment responder subpopulations. The population susceptible to HBV was also tracked by age and sex; it comprised uninfected individuals who had never been exposed to HBV and had not been successfully immunized.

The scheme of the modelled disease progression of HBV is shown in **Fig. A1**.

Newly infected cases entered the model through the incidence calculation described below. Those developing a chronic hepatitis B infection were split into low and high viral load cases using reported data on respective proportions of high viral load cases among populations negative or positive for hepatitis B e antigen (HBeAg). The risk for chronic hepatitis B infection largely depends on the age of acquisition of infection; hence, the model began in 1900 to allow for full flexibility of age of infection of the currently infected population, and to estimate the current susceptible population.

Fig. A1. Flow of the HBV disease progression model[a]



CHB: chronic hepatitis B; Cirr: compensated cirrhosis; DCC: decompensated cirrhosis; HCC: hepatocellular carcinoma; HVL: high viral load; LT: liver transplant; LVL: low viral load; T: treatment responder; U: untreated or non-responder.

[a] Black arrows = disease progression; orange arrows = non-liver-related death; red arrows = liver-related death; blue arrows = treatment response; purple arrows = treatment discontinuation; green arrows = liver transplantation.

Incidence

Annual incident cases of HBV infection by sex and age were calculated separately for perinatally and horizontally acquired infection (**Fig. A2**). Incident cases developing a chronic HBV infection were added to the disease progression model annually, and the resulting prevalence of HBsAg was used to calculate incident cases in the following year, generating a dynamic model (**Fig. A3**). Among those not developing a chronic infection, the risk for fulminant hepatitis B was 0.5% (UI: 0.1–1.0%), with the risk of death from fulminant hepatitis B being 67% (UI: 63–75%) for paediatrics and 90% (UI: 84–99%) for adults *(16, 17)*.

Fig. A2. Incidence determination scheme in the model

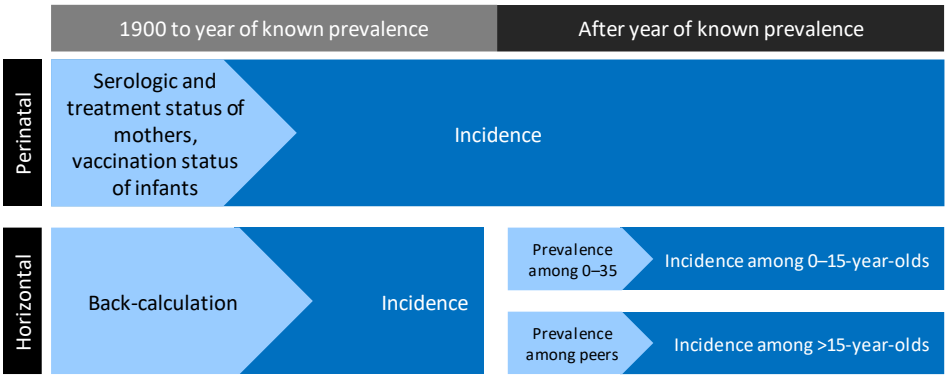
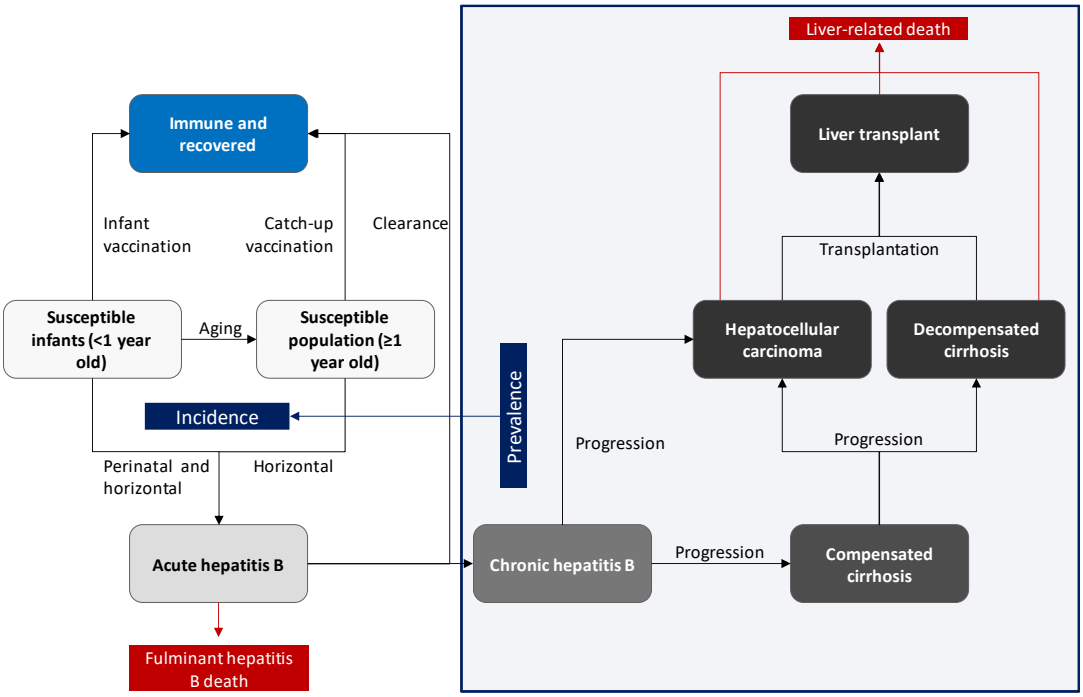


Fig. A3. Flow between populations in the model



Perinatal incidence

To calculate incident cases of perinatally acquired HBV infection, the annual modelled prevalence of HBsAg among women of childbearing age is subdivided into individuals who are estimated to be HBeAg-positive or HBeAg-negative. The proportions among these two groups that have a high viral load and low viral load are then combined into low viral load and high viral load groupings *(11, 13)*. The viral load grouping – in conjunction with the reported proportion of women with a high viral load receiving peripartum antiviral treatment – was used to segment the HBV-infected women of childbearing age into four groups according to serologic and treatment status: HBsAg-positive with high viral load untreated, HBsAg-positive with low viral load untreated, HBsAg-positive with high viral load treated and HBsAg-negative.

The reported number of annual births by mother's age group was then mapped to annual births by maternal serologic and treatment status, estimated by **Equation 1**.

Equation 1. Total births in year $\boldsymbol{t}$ to mothers of serologic and treatment status $\boldsymbol{i}$

$$\text{Total births}_{t,i} = \sum_{A} \text{Total births}_{t,A} \times g_{t,i,A}$$

where: A is the maternal age group (15–19, 20–24, …, 45–49)

$g_{t,i,A}$ is the proportion of A-year-old women of childbearing age at time t with serologic and treatment status i

Next, the births were segmented into five groups by vaccination status: no vaccination, birth dose of HBV vaccine only, complete HBV vaccine series with birth dose without HBIG, complete vaccine series with birth dose with HBIG, and complete vaccine series without birth dose.

Finally, using the transmission rates by the maternal serologic and treatment status and the infants' vaccination status (**Table A3**), the number of perinatally acquired cases of HBV infection was calculated (**Equation 2**).

Equation 2. New perinatally acquired cases of HBV at time $\boldsymbol{t}$

$$\text{New perinatally acquired cases of HBV}_{t} = \sum_{i} \left[\text{Total births}_{t,i} \times \sum_{j} \left(f_{i,j} \times r_{i,j} \right) \right]$$

where: i ranges over the maternal serologic and treatment status

j ranges over the vaccination status of infants

$f_{i,j}$ is the proportion of infants with vaccination status j born to mothers with serologic and treatment status i

$r_{i,j}$ is the transmission rate from a mother with serologic and treatment status i to an infant with vaccination status j

Uninfected infants who did not receive complete HBV vaccine series were added to the susceptible population. Chronically infected infants were added to the disease progression model.

Table A3. Mother-to-child transmission rates of HBV, percentage

Vaccination status of infant	Maternal serologic status		
	HBsAg-positive with high viral load, untreated peripartum	HBsAg-positive with low viral load, untreated peripartum	HBsAg-positive with high viral load, treated peripartum
No vaccination	100.0 (90.3–100.0)[11, 18-25]	0.0 (0.0–0.0)[11, 26]	5.6 (4.5–5.6)[a]
Birth dose of HBV vaccine only	90.0 (81.3–100.0)[18, 19, 27]	0.0 (0.0–0.0)[11, 26]	5.1 (4.1–5.6)[†]
Complete HBV vaccine series with birth dose without HBIG	13.8 (9.6–41.3)[27-29]	0.0 (0.0–0.0)[11, 26]	0.7 (0.5–2.1)[†]
Complete HBV vaccine series with birth dose with HBIG	7.7 (2.5–20.7)[26, 27, 30]	0.0 (0.0–0.0)[11, 26]	0.4 (0.1–1.0)[†]
Complete HBV vaccine series without birth dose	32.7 (29.0–35.0)[23]	0.0 (0.0–0.0)[11, 26]	1.6 (1.5–1.8)[†]

HBIG: hepatitis B immune globulin; HBsAg: hepatitis B surface antigen; HBV: hepatitis B virus.

[a] Assumed a 95% (response rate to treatment with tenofovir) reduction relative to transmission rates for HBsAg-positive with high viral load, untreated peripartum mothers

Horizontal incidence

Horizontally acquired incident cases of HBV infection were calculated separately, initially, up to the year of known prevalence by sex and age group in a country/region ('year of known prevalence'), and then after the year of known prevalence (**Fig. A2**). Cases incident up to the year of known prevalence are referred to as 'historical', whereas cases incident afterwards are referred to as 'forward'.

Horizontal incidence: historical

Total annual historical horizontally acquired incident cases of HBV infection were calculated by first defining a curve describing relative sizes of these incident cases ('relative incidence'). A calibration procedure matching modelled prevalence to reported prevalence was then used to transform relative incidence to annual incident cases by sex and age.

Relative incidence

Relative incidence was built by back-calculating arrays of quinquennial estimated incident cases for both sexes (**Table A4**) that satisfied the conditions in **Equations 3–4**. The results were then converted to annual estimated incident cases, linearly interpolated over 1900 to year of known prevalence, and passed through

a 5-year-average filter. The resulting annual incident cases were divided by peak incident cases to generate the relative incidence curve with a peak of one.

Table A4. Array of incident (horizontal) cases I_{t,s,A_n} at time t, of sex s, of age group A_n

Age group / Year	⋯	YKP − 5	YKP	YKP
A_0	⋯	$I_{\text{YKP}-5,s,A_0}$	I_{YKP,s,A_0}	P_{YKP,s,A_0}
A_1	⋯	$I_{\text{YKP}-5,s,A_1}$	I_{YKP,s,A_1}	P_{YKP,s,A_1}
A_2	⋯	$I_{\text{YKP}-5,s,A_2}$	I_{YKP,s,A_2}	P_{YKP,s,A_2}
⋮	⋱	⋮	⋮	⋮

YKP: year of known prevalence; P_{YKP,s,A_n}: total prevalent population in YKP of sex s in age group A_n; A_0 is age group 0, A_1 is age group 1–4, A_2 is age group 5–9, …, A_{18} is age group 85+.

Equation 3. Condition 1 for back-calculated array of incident (horizontal) cases

For each sex s and age group A_n,

$$\sum_{i=0}^{n} I_{\text{YKP}-5n+5i,s,A_i} \times \bar{c}_{A_i} \times S(d)_{\text{YKP}-5n+5i,s,A_i} \times \left(\text{HVL} \times S(l_{\text{HVL}})_{\text{YKP}-5n+5i,s,A_i} + (1-\text{HVL}) \times S(l_{\text{LVL}})_{\text{YKP}-5n+5i,s,A_i}\right) = P_{\text{YKP},s,A_n}$$

where:

YKP is year of known prevalence

I_{t,s,A_n} is incident (horizontal) cases at time t, of sex s, at age group A_n

A_0 is age group 0, A_1 is age group 1–4, A_2 is age group 5–9, …, A_{18} is age group 85+

$\bar{c}_{A_n}$ is average risk for chronic infection for age group A_n

$S(d)_{t,s,A_n}, S(l_{\text{HVL}})_{t,s,A_n}, S(l_{\text{LVL}})_{t,s,A_n}$ are survival functions from background death, liver-related death among high viral load population, and liver-related death among low viral load population, respectively, from time t to year of known prevalence for sex s, and age group A_n

HVL is proportion of high viral load cases among incident cases of HBV

P_{YKP,s,A_n} is prevalent cases of HBV, of sex s and age group A_n in year of known prevalence

Equation 4. Condition 2 for back-calculated array of incident (horizontal) cases

For each time point t, sex s, and age group A_n,

$$\frac{I_{t,s,A_n}}{I_{t-5,s,A_{n-1}}} = \frac{\bar{k}_{A_n}}{\bar{k}_{A_{n-1}}} \times \frac{\text{Unvaccinated population}_{t,s,A_n}}{\text{Unvaccinated population}_{t-5,s,A_{n-1}}}$$

where:

YKP is year of known prevalence

I_{t,s,A_n} is incident (horizontal) cases at time $t \in \{\text{YKP} - 85, \text{YKP} - 80, \ldots, \text{YKP}\}$, sex s, and age group $A_n \in \{A_1, \ldots, A_{18}\}$

$\bar{k}_{A_n}$ is average shape parameter for age group A_n

$\text{Unvaccinated population}_{t,s,A_n}$ at time t, of sex s, at age group A_n was estimated using reported history of vaccinations in the country

Incidence calibration

A (1) scalar multiplier of relative incidence and (2) quinquennial sex and age group distributions of historical horizontal incident cases were calculated using the secant method *(31)* to match (1) modelled total prevalent cases (**Equation 5**) and (2) modelled prevalence of HBV infection by sex and age group in the year of the known prevalence to reported prevalence.

Equation 5. Total HBV-infected population in year of known prevalence

$$\text{Prevalent cases of HBV infection}_{\text{YKP}} = \sum_{t=1900}^{\text{YKP}} \left(\text{Incident cases of chronic HBV infection}_t - \text{Deaths among HBV-infected population}_t\right)$$

where:

YKP is year of known prevalence

Horizontal incidence: forward

Susceptible population

To calculate the number of new horizontally acquired HBV infections after the year of known prevalence, the first step was to estimate the susceptible population in the year of known prevalence (**Equation 6**). Numbers of susceptible infants were calculated as described above.

Equation 6. Susceptible population S in year of known prevalence, of sex s, at age $a \geq 1$

$$S_{\text{YKP},s,a} = \left(P_{\text{YKP},s,a} - \sum_{i=0}^{a} \left[\text{Incident HBV cases}_{\text{YKP}-a+i,s,i} \prod_{j=i}^{a-1} \left(1 - d_{\text{YKP}-a+j,s,j}\right) \right] \right) \times \left(1 - \text{Imm}_{\text{YKP},a}\right)$$

where:

YKP is year of known prevalence

$d_{t,s,a}$ is background mortality rate at time t, of sex s, at age a

$P_{t,s,a}$ is population at time t, of sex s, at age a

$\text{Imm}_{t,a}$ is an estimate of immunization coverage of population at time t at age a

In all years following the year of known prevalence, the annual susceptible population was calculated by sex and age by first adding infants susceptible to infection to the existing 0-year-old susceptible population, as described above, and then subtracting deaths due to background mortality, new infections and new catch-up immunizations from the existing susceptible population (**Equation 7**).

Equation 7. Size of susceptible population S after year of known prevalence, of sex s, at age $a \geq 1$

$$S_{t,s,a} = S_{t-1,s,a-1} \times \left(1 - d_{t-1,s,a-1}\right) - \text{New HBV cases}_{t,s,a} - \text{Catch-up immunizations}_{t,s,a}$$

where:

$S_{t-1,s,a-1}$ is the size of the susceptible population at time $t-1$, of sex s, at age $a-1$

$d_{t,s,a}$ is background mortality rate at time t, of sex s, at age a

Incidence function

The incidence of horizontally acquired HBV infection was assumed to be a linear function of HBV prevalence with high viral load (**Equation 8**). For those aged under 15 years, incidence was determined by prevalence of HBV infection with high viral load among those aged 0–35 years to simulate household infection from siblings, peers, parents and other adults. For those aged 15 years or older, incidence was a function of prevalence with high viral load among peers of the same age.

Equation 8. Number of incident (horizontal) cases I in year $t >$ YKP, of sex s, and age a

$$I_{t,s,a} = \begin{cases} S_{t-1,s,a-1} \times \left(1 - d_{t-1,s,a-1}\right) \times p_{t-1,0\text{–}35} \times (k_a \times C_s) & \text{if } 0 \leq a < 15 \\ S_{t-1,s,a-1} \times \left(1 - d_{t-1,s,a-1}\right) \times p_{t-1,a} \times (k_a \times C_s) & \text{if } a \geq 15 \end{cases}$$

where:

$p_{t,a}$ is prevalence of HBV infection with high viral load at time t, at age (group) a

$S_{t,s,a}$ is the size of the susceptible population at time t, of sex s, at age a

$d_{t,s,a}$ is background mortality rate at time t, of sex s, at age a

k_a is the shape parameter for age a

C_s is scale parameter for sex s (see below)

Scale parameter C_s was inferred through a calibration procedure for both sexes with an s variable that matched the number of cases incident in the year of known prevalence to those incident in the following year; this approach tied the forward incident cases to historical incident cases. Shape parameter k_a for each age a was assumed to be constant across all countries.

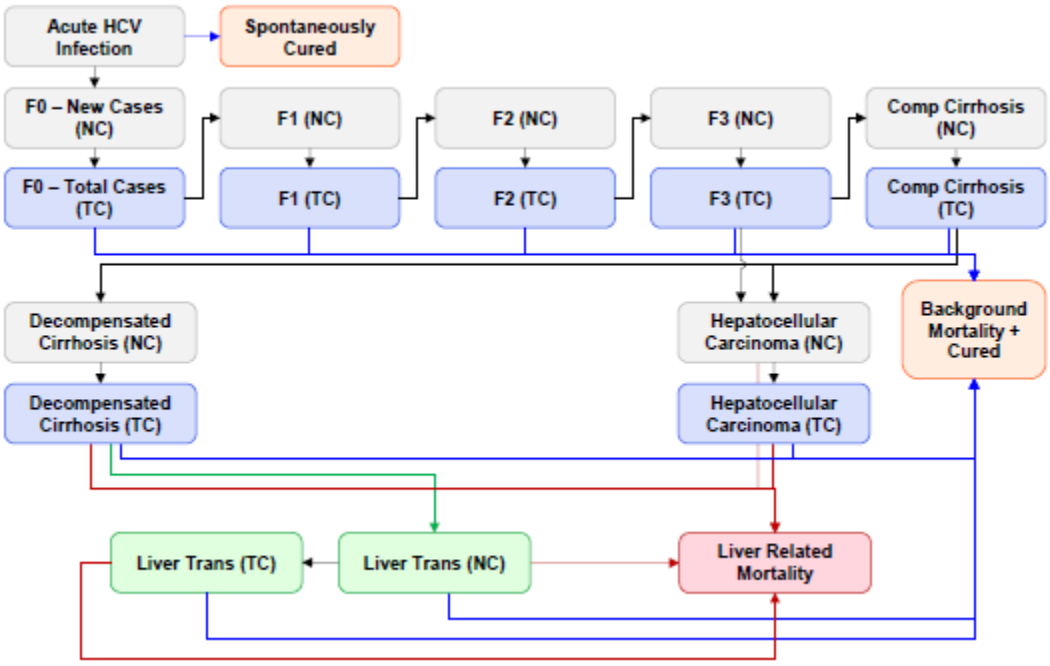
4.3 HCV modelling methods and inputs

The HCV disease burden model used in this report builds on the modelling framework developed by the Polaris Observatory; the model has been refined through multiple iterations and expert reviews since its inception in 2012. The HCV model is a compartmental, deterministic Markov model constructed in MS Excel. As with HBV, MS Excel was selected for transparency, accessibility and ease of review by national experts. The model quantifies the annual number of viraemic HCV infections and tracks disease progression, stratified by age, sex and stage of liver disease, from 1950 to 2050.

The model follows individuals with chronic HCV infection (HCV RNA positive) and estimates transitions across key disease stages: F0, F1, F2, F3, compensated cirrhosis, decompensated cirrhosis, HCC and liver transplantation. An all-cause mortality is applied to cases in each stage of liver disease, and an additional liver-related mortality is applied to cases in end-stage liver disease and in people who have received a liver transplant. Acute infections that spontaneously clear (i.e. non-viraemic) are not carried forward, because the model focuses on the viraemic population burden.

The structure of the model, including the progression pathways, is presented in **Fig. A4**.

Fig. A4. The flow of the HCV disease progression model



Cirr: compensated cirrhosis; DCC: decompensated cirrhosis; HCC: hepatocellular carcinoma; HCV: hepatitis C virus; LT: liver transplant; NC: new cases; TC: total cases.

Model structure and logic

New HCV infections enter the model through two routes – **vertically acquired infections** and **horizontally acquired infections –** as described below. Individuals who develop chronic infection are assigned to the appropriate **fibrosis stage (F0)**, and disease progression thereafter is determined by **age- and sex-specific transition rates**, using published meta-analyses and adjusted to reflect empirical HCC incidence patterns in high-quality national datasets.

Each year, transitions are based on:

- fibrosis progression rates (F0→F1→F2→F3→cirrhosis);
- progression to decompensation and HCC;
- liver transplantation rates;
- liver-related and background mortality;
- treatment initiation; and
- cure (SVR).

Cured individuals exit the viraemic pool; those who fail treatment remain in their previous stage.

The general equation governing annual prevalence at each stage is provided in **Equations 9–11**.

Disease stages and transitions: the HCV model tracks the following stages:

- acute infection → spontaneous clearance or chronic infection (viraemic);
- chronic infection, F0–F3;
- compensated cirrhosis;

- decompensated cirrhosis;
- HCC;
- liver transplant; and
- death (background and liver-related).

Progression rates are **back-calculated** using 5-year age cohorts and applying published estimates; the rates are then refined using HCC age-specific incidence in the USA, after adjusting for the proportion attributable to HCV infection.

Incidence calculation

Vertical transmission: Vertical infections are estimated using age-specific HCV prevalence among women of childbearing age, fertility rates and a standard HCV mother-to-child transmission rate of 5.8% (95% confidence interval [CI]: 4.2–7.8%) *(32)*. Although transmission is higher among HIV-positive women, the global HIV/HCV coinfection prevalence is low; therefore, vertical transmission estimates were not adjusted for HIV.

Horizontal transmission: Historical incident cases are calibrated to reported overall, as well as sex- and age-group-specific prevalence of chronic HCV infection in a country or territory, using the calibration procedure described below:

> **Historical incidence** (1950 → year of known prevalence): historical incidence is reconstructed using two strategies:
>
> - Two data points available: when there are two prevalence studies with age- and sex-specific distributions, the model calculates the average annual incidence required to move from the earlier to the later prevalence, adjusting for mortality and past treatment.
> - One data point available: when there is only one reliable prevalence study, a back-calculation approach is applied using assumed negligible prevalence in 1950, mortality and cure patterns, and a relative incidence curve reflecting infection trends (e.g. nosocomial peaks, blood-screening introduction and waves driven by people who inject drugs). This curve is mapped using country-specific risk factor timelines and validated with national experts.
>
> **Future incidence (after the last data year)**: this is assumed to change at the same annual rate as chronic prevalence. In settings with fibrosis-restricted treatment eligibility (e.g. ≥F2 policies), the trend follows **F0 prevalence**; otherwise, it follows **overall prevalence**.

Prevalence by age: Prevalence data are often incomplete across all age cohorts. The model applies the following rules: older ages (>50 years) – prevalence held constant at the oldest measured cohort; younger ages (<18 years) – exponential decline applied; missing intermediate cohorts – smoothed using epidemiological plausibility checks (e.g. Malaysia and Burundi examples). These assumptions ensure age-specific prevalence curves required for Markov initialization.

Treatment, cure and cascade dynamics: Annual treatment numbers come from national registries, drug-sales audits, programme reports and expert confirmation. Treatment allocation follows treatment eligibility by fibrosis stage and age**,** distribution of treated patients across genotypes**,** real-world SVR rates (for interferon, triple therapy and DAA-only regimens), and forecasted decline from treatment peaks (logarithmic decay over 5 years unless country-specific data override). Cured cases (SVR) are removed from the viraemic population annually.

Diagnosis and screening: Diagnosis estimates rely on national surveillance, registries or regional analogues, adjusted for spontaneous clearance, viraemic proportion of diagnosed individuals, and background screening intensity. A screening module is incorporated to estimate the **number needed to screen (NNS)** to detect anti-HCV cases based on undiagnosed prevalence, linkage-to-care status, age-based or cohort-based screening policies, and the permitted number of lifetime screens.

Calibration and model fitting: The model is calibrated to match total HCV prevalence in the year of known prevalence, and age- and sex-specific prevalence profiles; the calibration considers empirical HCC incidence (where available), and historical diagnosis and treatment trends. A secant method is used to identify the constant that scales the relative incidence curve, to yield observed prevalence after adjusting for mortality and cure.

Outputs are validated using observed trends from longitudinal prevalence studies (e.g. France, Egypt and the USA), HCC incidence data (e.g. GLOBOCAN adjusted for HCV attribution) and expert consensus through the Delphi process.

Equation 9. Total HCV cases

$$\begin{aligned}\text{Total cases}_{\text{Stage}_x,\,\text{Year}_y,\,\text{Age}_a} &= \text{Total cases}_{\text{Stage}_x,\,\text{Year}_{y-1},\,\text{Age}_{a-1}} + \text{New cases}_{\text{Stage}_x,\,\text{Year}_y,\,\text{Age}_a} - \text{Cured}_{\text{Stage}_x,\,\text{Year}_y,\,\text{Age}_a} \\ &- \text{Background mortality}_{\text{Stage}_x,\,\text{Year}_y,\,\text{Age}_a} - \text{Progressed}_{\text{Stage}_x,\,\text{Year}_y,\,\text{Age}_a} \\ &- \text{Liver-related mortality}_{\text{Stage}_x,\,\text{Year}_y,\,\text{Age}_a}\end{aligned}$$

Equation 10. New HCV cases

$$\text{New cases}_{\text{Stage}_x,\,\text{Year}_y,\,\text{Age}_a} = (\text{Total cases}_{\text{Stage}_{x-1},\,\text{Year}_{y-1},\,\text{Age}_{a-1}}) \times (\text{Progression rate}_{\text{Stage}_{x-1}\rightarrow\text{Stage}_x,\,\text{Age}_{a-1}})$$

Equation 11. HCV cured cases

$$\begin{aligned}\text{Cured}_{\text{Stage}_x,\,\text{Year}_y,\,\text{Age}_a} &= (\text{Total cases}_{\text{Stage}_x,\,\text{Year}_{y-1},\,\text{Age}_{a-1}}) \times (\text{Age-eligibility flag}_{\text{Year}_{y-1},\,\text{Age}_{a-1}}) \\ &\times \left(\frac{\text{Cured}_{\text{Stage}_x,\,\text{Year}_y}}{\text{Total age-eligible cases}_{\text{Stage}_x,\,\text{Year}_{y-1}}}\right)\end{aligned}$$

Where:

$$\text{Cured}_{\text{Stage}_x,\,\text{Year}_y} = \sum_{g=1}^{6} (\text{Total treated}_{g,\,\text{Stage}_x,\,\text{Year}_y} \times \text{SVR}_{g,\,\text{Year}_y})$$

Equation 12. Background mortality

$$\begin{aligned}\text{Background mortality}_{\text{Stage}_x,\,\text{Year}_y,\,\text{Age}_a} &= (\text{Total cases}_{\text{Stage}_x,\,\text{Year}_{y-1},\,\text{Age}_{a-1}} \\ &- \text{Cured}_{\text{Stage}_x,\,\text{Year}_y,\,\text{Age}_a}) \times (\text{Adjusted background mortality rate}_{\text{Year}_{y-1},\,\text{Age}_{a-1}})\end{aligned}$$

Equation 13. **Progressed HCV cases**

$$\begin{aligned}\text{Progressed}_{\text{Stage}_x,\,\text{Year}_y,\,\text{Age}_a} &= (\text{Total cases}_{\text{Stage}_x,\,\text{Year}_{y-1},\,\text{Age}_{a-1}} - \text{Cured}_{\text{Stage}_x,\,\text{Year}_y,\,\text{Age}_a} \\ &\quad - \text{Background mortality}_{\text{Stage}_x,\,\text{Year}_y,\,\text{Age}_a}) \times \text{Progression rate}_{\text{Stage}_x \to \text{Stage}_{x+1},\,\text{Age}_{a-1}}\end{aligned}$$

Equation 14. **Liver-related mortality**

$$\begin{aligned}\text{Liver-related mortality}_{\text{Stage}_x,\,\text{Year}_y,\,\text{Age}_a} &= (\text{Total cases}_{\text{Stage}_x,\,\text{Year}_{y-1},\,\text{Age}_{a-1}} - \text{Cured}_{\text{Stage}_x,\,\text{Year}_y,\,\text{Age}_a} \\ &\quad - \text{Background mortality}_{\text{Stage}_x,\,\text{Year}_y,\,\text{Age}_a} \\ &\quad - \text{Progressed}_{\text{Stage}_x,\,\text{Year}_y,\,\text{Age}_a}) \times \text{Liver-related mortality rate}_{\text{Year}_{y-1},\,\text{Age}_{a-1}}\end{aligned}$$

5. HBV model validation

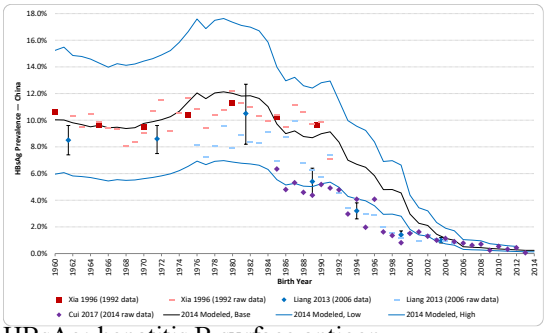
The primary objective of the PRoGReSs model is to accurately predict HBsAg prevalence by age and sex. During the literature review, prevalence studies that had at least two time points were examined for their viability in testing the model outputs against literature-reported data. Three countries were chosen: China, the Islamic Republic of Iran and Uganda. These countries differ in HBsAg prevalence, HBeAg prevalence and genotype. Although the model's ability to predict future prevalence was examined for each of these countries, most recent data were used in the models that were used for the national, regional and global estimates.

China

For China, prevalence data were available by 5-year or 10-year age groups at two time points (1992 and 2006) and by birth year at three time points (1992, 2006 and 2014) *(33-35)*. Two rounds of validation were conducted.

In the first round of validation, the model was calibrated to 1992 prevalence data *(34)*. Using the range provided in those data, along with the aforementioned uncertainty in model inputs, a probabilistic uncertainty analysis was conducted. The modelled prevalence in 2014 by birth year was compared with the reported data from 1992, 2006 and 2014 (**Fig. A5**). Only the 2006 data were provided with CIs for specific age groups. In 2014, the model was able to predict prevalence within the CI of 2006 data for those born before 1989. However, there was a large drop in prevalence among the 1989 birth cohort compared with the 1990 birth cohort in the 1992 data. Owing to this discrepancy in reported prevalence, the baseline model had difficulty accurately predicting prevalence for birth year cohorts from 1989 to the end of 2004. Nevertheless, the low output of the probabilistic sensitivity analysis accurately predicted the 2014 prevalence by birth year – with the exception of a few birth years in the late 1990s.

Fig. A5. Validation of modelled outcomes – China, round 1

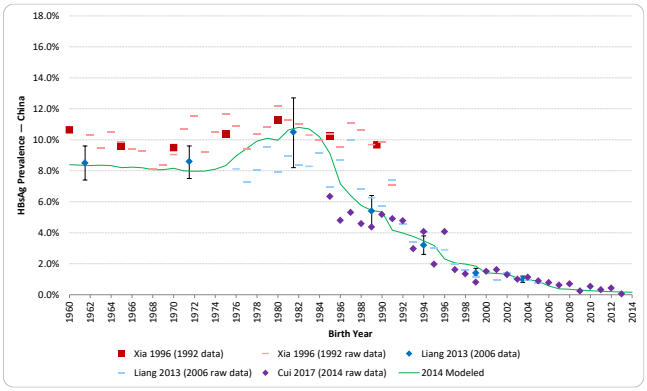


HBsAg: hepatitis B surface antigen.

In the second round of validation, the model was calibrated to 2006 prevalence data *(33)*. Modelled prevalence by year was compared with all available prevalence data from the three time points (**Fig. A6**). When calibrated to the 2006 data, the model accurately predicted the prevalence that was reported in 2014, particularly among those born after the year of calibration. The validation rounds for data from China

provide evidence that the model can accurately predict age-specific prevalence, depending on the quality of the epidemiological inputs.

Fig. A6. Validation of modelled outcomes – China, round 2

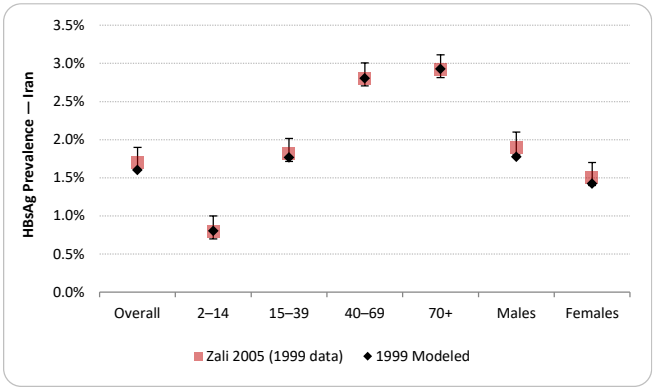


HBsAg: hepatitis B surface antigen.

The Islamic Republic of **Iran**

Before the introduction of national HBV vaccination in the Islamic Republic of Iran, a serologic survey was conducted in 1990 and again in 1999 *(36)*. For this validation, the model was calibrated to the 1990 data and the modelled outputs were compared with reported prevalence in 1999 (**Fig. A7**). The model predicted prevalence reported by age group quite well, and the reported overall prevalence and prevalence by sex were within the CIs. Part of the observed discrepancy was because the 1990 survey did not sample people aged 70 years or older.

Fig. A7. Validation of modelled outcomes – the Islamic Republic of Iran



HBsAg: hepatitis B surface antigen.

Uganda

The model was calibrated to available data from Uganda, which had a national serologic survey in 2005 that only included adults (**Fig. A8**) *(37)*. This was used in the model to estimate the prevalence by age in 1990 (**Fig. A9**).

Fig. A8. Validation of modelled outcomes (2005) – Uganda

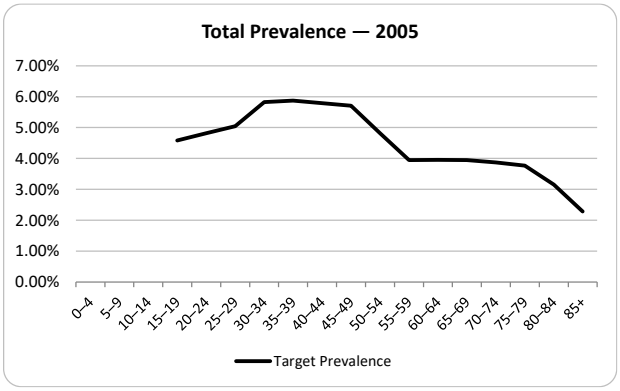
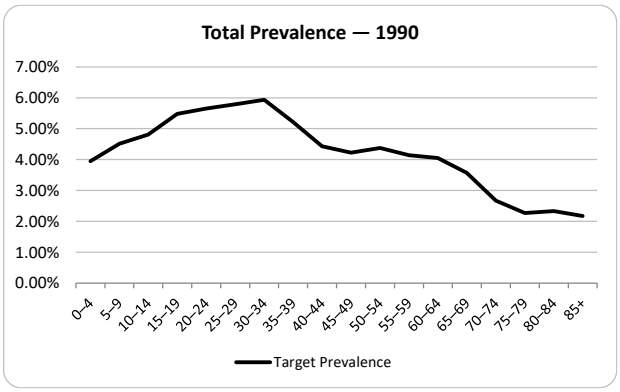
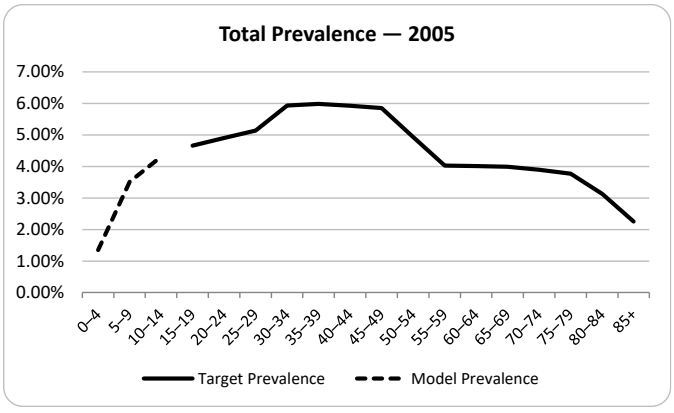


Fig. A9. Validation of modelled outcomes (1990) – Uganda



The model was then calibrated to the 1990 modelled prevalence data and run forward to 2005, to estimate the prevalence in the paediatric age groups that were not included in the national serosurvey, accounting for transmission (both mother-to-child and horizontal), prophylaxes measures and mortality that occurred over the period. This was of particular importance because national vaccination began in 2002. The results of this exercise, as well as the impact of vaccination on the 0–4-year-old cohort, are shown in **Fig. A10**.

Fig. A10. Validation of modelled outcomes (2005) – Uganda



A new national serologic survey, which was published in 2019, estimated that in 2017 the prevalence among 0–14-year-old boys was 0.7%, and 0.6% among girls *(38)*. These data were compared with the modelled outputs, after estimating paediatric prevalence since 1991 (a timeframe of 26 years), with the model estimating HBV prevalence of 0.7% and 0.5% among boys and girls, respectively, for the group aged 0–14 years.

These three countries provided evidence that, with reliable input data, the model can accurately predict sex- and age-specific as well as overall prevalence over many years.

6. Limitations of modelled estimates

- The estimates presented may differ from official national figures owing to the use of different data sources. A common discrepancy occurs when the model relies on a single historical data point (e.g. a serosurvey from 2014) and projects it forward to the reporting year. In such cases, the model incorporates the expected effects of prophylaxis, transmission and mortality over the intervening years.
- It is well understood that data obtained by literature review, input from experts, external sources and modelling will not entirely capture the reality of the field. For these reasons, a validation step at the regional and national level through the WHO country consultation process is crucial in order to reach consensual agreement.
- Although these data provide the most recent estimates based on currently available evidence in countries, they might obscure regional variations and not account for certain populations at high risk (e.g. migrants, Indigenous peoples and nations) that might have an increased prevalence.
- Although these are the latest national-level estimates, they do not explicitly assess differences between rural and urban settings, where access to prophylaxis may vary. These limitations highlight the need for national strategies tailored to the regions and populations most affected by HBV infection.
- The impacts of immigration and emigration are not included in these calculations. As a result, prevalence may be underestimated in low-prevalence countries receiving migrants from high-prevalence settings.
- Extrapolated estimates for children aged 5 years or below may underestimate or overestimate the actual prevalence in countries lacking direct data.

References

1. Blach S, Zeuzem S, Manns M, Altraif I, Duberg A-S, Muljono DH et al. Global prevalence and genotype distribution of hepatitis C virus infection in 2015: a modelling study. Lancet Gastroenterol Hepatol. 2017;2(3):161–76 (https://doi.org/10.1016/S2468-1253(16)30181-9).
2. Blach S, Terrault NA, Tacke F, Gamkrelidze I, Craxi A, Tanaka J et al. Global change in hepatitis C virus prevalence and cascade of care between 2015 and 2020: a modelling study. Lancet Gastroenterol Hepatol. 2022;7(5):396–415 (https://doi.org/10.1016/S2468-1253(21)00472-6).
3. Razavi-Shearer D, Gamkrelidze I, Pan C, Jia J, Berg T, Gray R et al. Global prevalence, cascade of care, and prophylaxis coverage of hepatitis B in 2022: a modelling study. Lancet Gastroenterol Hepatol. 2023;8(10):879–907 (https://doi.org/10.1016/S2468-1253(23)00197-8).
4. Razavi-Shearer D, Gamkrelidze I, Nguyen MH, Chen D-S, Van Damme P, Abbas Z et al. Global prevalence, treatment, and prevention of hepatitis B virus infection in 2016: a modelling study. Lancet Gastroenterol Hepatol. 2018;3(6):383–403 (https://doi.org/10.1016/S2468-1253(18)30056-6).
5. Polaris Observatory [website]. CDA Foundation; 2026 (https://cdafound.org/polaris).
6. Polaris Observatory Collaborators. Global prevalence, cascade of care, and prophylaxis coverage of hepatitis B in 2022: a modelling study. 2023 (
7. Razavi HA, Waked I, Qureshi H, Kondili LA, Duberg AS, Aleman S et al. Number of people treated for hepatitis C virus infection in 2014–2023 and applicable lessons for new HBV and HDV therapies. J Hepatol. 2025 (https://doi.org/10.1016/j.jhep.2025.01.013).
8. World Population Prospects 2024: summary of results (UN DESA/POP/2024/TR/NO. 9).Department of Economic and Social Affairs, Population Division, United Nations; 2024 (https://population.un.org/wpp/).
9. Marcellin P, Gane E, Buti M, Afdhal N, Sievert W, Jacobson IM et al. Regression of cirrhosis during treatment with tenofovir disoproxil fumarate for chronic hepatitis B: a 5-year open-label follow-up study. Lancet. 2013;381(9865):468–75 (https://doi.org/10.1016/S0140-6736(12)61425-1).
10. World Health Organization, United Nations Children's Fund. WHO/UNICEF estimates of national immunization coverage (WUENIC), July 2025 release.WHO; 2025 (https://www.who.int/teams/immunization-vaccines-and-biologicals/immunization-analysis-and-insights/global-monitoring/immunization-coverage/who-unicef-estimates-of-national-immunization-coverage).
11. Wiseman E, Fraser MA, Holden S, Glass A, Kidson BL, Heron LG et al. Perinatal transmission of hepatitis B virus: an Australian experience. Med J Aust. 2009;190(9):489–92 (https://www.ncbi.nlm.nih.gov/pubmed/19413519).
12. Chen CJ, Yang HI. Natural history of chronic hepatitis B REVEALed. J Gastroenterol Hepatol. 2011;26(4):628–38 (https://doi.org/10.1111/j.1440-1746.2011.06695.x).
13. Szymańska K, Chen JG, Cui Y, Gong YY, Turner PC, Villar S et al. TP53 R249S mutations, exposure to aflatoxin, and occurrence of hepatocellular carcinoma in a cohort of chronic hepatitis B virus carriers from Qidong, China. Cancer Epidemiol Biomarkers Prev. 2009;18(5):1638–43 (https://doi.org/10.1158/1055-9965.epi-08-1102).
14. Edmunds W, Medley G, Nokes D, Hall A, Whittle H. The influence of age on the development of the hepatitis B carrier state. Proc R Soc B. 1993;253(1337):197–201 (https://doi.org/10.1098/rspb.1993.0102).
15. IRODaT: International Registry in Organ Donation and Transplantation [website]. IRODaT; 2026 (https://www.irodat.org/).
16. Davis GL, Alter MJ, El-Serag H, Poynard T, Jennings LW. Aging of hepatitis C virus (HCV)-infected persons in the United States: a multiple cohort model of HCV prevalence and disease progression. Gastroenterology. 2010;138(2):513–21.e6 (https://doi.org/10.1053/j.gastro.2009.09.067).
17. Chang MH. Hepatitis B virus infection. Semin Fetal Neonatal Med. 2007;12(3):160–7 (https://doi.org/10.1016/j.siny.2007.01.013).
18. Xu ZY, Liu CB, Francis DP, Purcell RH, Gun ZL, Duan SC et al. Prevention of perinatal acquisition of hepatitis B virus carriage using vaccine: preliminary report of a randomized, double-blind placebo-controlled and comparative trial. Pediatrics. 1985;76(5):713–8 (https://www.ncbi.nlm.nih.gov/pubmed/3903646).
19. Okada K, Kamiyama I, Inomata M, Imai M, Miyakawa Y. e antigen and anti-e in the serum of asymptomatic carrier mothers as indicators of positive and negative transmission of hepatitis B virus to their infants. N Engl J Med. 1976;294(14):746–9 (https://doi.org/10.1056/nejm197604012941402).
20. Luo Z, Li L, Ruan B. Impact of the implementation of a vaccination strategy on hepatitis B virus infections in China over a 20-year period. Int J infect Dis. 2012;16(2):e82–e8 (https://doi.org/10.1016/j.ijid.2011.10.009).

21. Beasley RP, Hwang LY, Lin CC, Ko YC, Twu SJ. Incidence of hepatitis among students at a university in Taiwan. Am J Epidemiol. 1983;117(2):213–22 (https://doi.org/10.1093/oxfordjournals.aje.a113532).
22. Ip HM, Lelie PN, Wong VC, Kuhns MC, Reesink HW. Prevention of hepatitis B virus carrier state in infants according to maternal serum levels of HBV DNA. Lancet. 1989;1(8635):406–10 (https://www.ncbi.nlm.nih.gov/pubmed/2563788).
23. Lo KJ, Tsai YT, Lee SD, Yeh CL, Wang JY, Chiang BN et al. Combined passive and active immunization for interruption of perinatal transmission of hepatitis B virus in Taiwan. Hepatogastroenterology. 1985;32(2):65–8 (PMID:3159639).
24. Nayak NC, Panda SK, Zuckerman AJ, Bhan MK, Guha DK. Dynamics and impact of perinatal transmission of hepatitis B virus in North India. J Med Virol. 1987;21(2):137–45 (https://doi.org/10.1002/jmv.1890210205Digital).
25. Wong VC, Ip HM, Reesink HW, Lelie PN, Reerink-Brongers EE, Yeung CY et al. Prevention of the HBsAg carrier state in newborn infants of mothers who are chronic carriers of HBsAg and HBeAg by administration of hepatitis B vaccine and hepatitis B immunoglobulin. Double-blind randomised placebo-controlled study. Lancet. 1984;1(8383):921–6 (https://doi.org/10.1016/s0140-6736(84)92388-2).
26. Zou H, Chen Y, Duan Z, Zhang H, Pan C. Virologic factors associated with failure to passive–active immunoprophylaxis in infants born to HBsAg-positive mothers. J Viral Hepat. 2012;19(2):e18–e25 (https://doi.org/10.1111/j.1365-2893.2011.01492.x).
27. Poovorawan Y, Sanpavat S, Pongpunlert W, Chumdermpadetsuk S, Sentrakul P, Safary A. Protective efficacy of a recombinant DNA hepatitis B vaccine in neonates of HBe antigen-positive mothers. JAMA. 1989;261(22):3278–81 (https://doi.org/10.1001/jama.1989.03420220092033).
28. Kang G, Ma F, Chen H, Yang Y, Guo S, Wang Z et al. Efficacy of antigen dosage on the hepatitis B vaccine response in infants born to hepatitis B-uninfected and hepatitis B-infected mothers. Vaccine. 2015;33(33):4093–9 (https://doi.org/10.1016/j.vaccine.2015.06.081).
29. Andre FE, Zuckerman AJ. Review: protective efficacy of hepatitis B vaccines in neonates. J Med Virol. 1994;44(2):144–51 (https://www.ncbi.nlm.nih.gov/pubmed/7852954).
30. Kubo A, Shlager L, Marks AR, Lakritz D, Beaumont C, Gabellini K et al. Prevention of vertical transmission of hepatitis B: an observational study. Ann Intern Med. 2014;160(12):828–35 (https://doi.org/10.7326/M13-2529).
31. Weisstein EW. Secant method [website]. Wolfram Research, Inc.; 2026 (http://mathworld.wolfram.com/SecantMethod.html).
32. Benova L, Mohamoud YA, Calvert C, Abu-Raddad LJ. Vertical transmission of hepatitis C virus: systematic review and meta-analysis. Clin Infect Dis. 2014;59(6):765–73 (https://doi.org/10.1093/cid/ciu447).
33. Liang X, Bi S, Yang W, Wang L, Cui G, Cui F et al. Reprint of: Epidemiological serosurvey of Hepatitis B in China-Declining HBV prevalence due to Hepatitis B vaccination. Vaccine. 2013;31(S9):J21–J8 (https://doi.org/10.1016/j.vaccine.2013.08.012).
34. Xia GL, Liua CB, Caoa HL, Bia SL, Zhana MY, Sub CA et al. Prevalence of hepatitis B and C virus infections in the general Chinese population. Results from a nationwide cross-sectional seroepidemiologic study of hepatitis A, B, C, D, and E virus infections in China, 1992. Hepatol Commun. 1996;5(1):62–73 (https://doi.org/10.1016/S0928-4346(96)82012-3).
35. Wang F, Zhang G, Zheng H, Miao N, Shen L, Wang F et al. Post-vaccination serologic testing of infants born to hepatitis B surface antigen positive mothers in 4 provinces of China. Vaccine. 2017;35(33):4229–35 (https://doi.org/10.1016/j.vaccine.2017.06.019).
36. Zali MR, Mohammad K, Noorbala AA, Noorimayer B, Shahraz S. Rate of hepatitis B seropositivity following mass vaccination in the Islamic Republic of Iran. East Mediterr Health J. 2005;11(1-2):62–7 (PMID:16532672).
37. Bwogi J, Braka F, Makumbi I, Mishra V, Bakamutumaho B, Nanyunja M et al. Hepatitis B infection is highly endemic in Uganda: findings from a national serosurvey. Afr Health Sci. 2009;9(2):98–108 (PMID: 19652743).
38. Uganda Population-based HIV Impact Assessment (UPHIA) 2016–2017: final report.Ministry of Health Uganda; 2019 (https://phia.icap.columbia.edu/wp-content/uploads/2020/02/UPHIA_Final_Report_Revise_07.11.2019_Final_for-web.pdf).