*Big Data and Information Analytics*

https://www.aimspress.com/journal/BDIA

BDIA, 10: 294–307.
DOI:10.3934/bdia.2026015
Received: 28 September 2026
Accepted: 05 October 2026
Published: 27 October 2026

*Research article*

# Hierarchical Bayesian modelling of HBV DNA decay: comparison of viral decline following telbivudine and tenofovir disoproxil fumarate prophylaxis

**Dayaram Nehra**[1] **and Daihai He**[2,*]

[1] Department of Mathematics, Indian Institute of Science Education and Research Mohali, India

[2] Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong SAR, China

* **Correspondence:** Email: daihai.he@polyu.edu.hk.

**Abstract:** Antiviral prophylaxis during pregnancy can lower maternal hepatitis B virus (HBV) DNA and thus help reduce the risk of mother-to-child transmission. However, the decline in the viral load can differ substantially between patients and may also change as treatment continues. We analyzed longitudinal HBV DNA measurements from 328 pregnant women who received either telbivudine (LdT; $n = 95$) or tenofovir disoproxil fumarate (TDF; $n = 233$) with a hierarchical Bayesian model. The treatment comparison was observational. Because several potentially important clinical characteristics were not available for linkage to the individual trajectories, the adjusted analysis used the last HBV DNA measurement at or before treatment initiation and the treatment initiation week as patient-level covariates. We modeled the decline with a continuous two-phase exponential curve, allowing a separate rate during the first 12 weeks and after week 12. In the adjusted analysis, the posterior mean early-phase decline rates were 0.0964 week$^{-1}$ for LdT and 0.0752 week$^{-1}$ for TDF, a difference of 0.0212 week$^{-1}$ (95% CrI: 0.0100–0.0332). After week 12, the corresponding rates were 0.0084 and 0.0143 week$^{-1}$, with a difference of −0.0059 week$^{-1}$ (95% CrI: −0.0132–0.0021). The model-based population mean time to 5.3 $\log_{10}$ IU/mL was 3.75 weeks for LdT and 4.56 weeks for TDF. When the specified low-end observations were excluded, the early-rate difference fell to 0.0048 week$^{-1}$ (95% CrI: −0.0029–0.0131). Overall, both groups showed a marked early decline followed by a slower later phase. The apparent early treatment difference was sensitive to the handling of low-end measurements, and the observational design does not permit a causal treatment conclusion. Nevertheless, the Bayesian framework provides a useful way to describe patient-to-patient variations and uncertainty in these longitudinal HBV DNA trajectories.

**Keywords:** hepatitis B virus; hierarchical Bayesian model; viral-load dynamics; antiviral prophylaxis; longitudinal data; telbivudine; tenofovir disoproxil fumarate; posterior predictive assessment

## 1. Introduction

Chronic hepatitis B virus (HBV) infection remains an important global health problem, and maternal infection is particularly relevant because high maternal viral load is associated with an increased risk of mother-to-child transmission. The European Association for the Study of the Liver's 2017 clinical practice guidelines address the management of HBV infection [1]. The World Health Organization's 2024 guideline provides recommendations for the prevention, diagnosis, care, and treatment of chronic HBV infection [2].

Pan et al. evaluated tenofovir disoproxil fumarate (TDF) for preventing HBV transmission in mothers with high viral loads [3]. Brown et al. conducted a systematic review and meta-analysis of antiviral therapy for chronic HBV infection during pregnancy [4]. Together, these studies provide clinical context for examining viral-load changes during antiviral prophylaxis.

Nowak et al. developed mathematical descriptions of viral dynamics in HBV infection [5]. Tsiang et al. reported biphasic HBV clearance kinetics during adefovir dipivoxil therapy [6], while Wolters et al. studied viral dynamics in chronic HBV patients receiving lamivudine [7]. Perelson and Ribeiro reviewed HBV kinetics and mathematical modelling [8]. Ciupe subsequently reviewed models of HBV infection, immunity, and drug therapy [9]. These works motivate models that represent longitudinal viral-load changes while recognizing that different formulations capture different aspects of the underlying process.

More recently, Shen et al. [10] analyzed real-world longitudinal HBV DNA data from 328 pregnant women who received either telbivudine (LdT) or tenofovir disoproxil fumarate (TDF). Their study used a sigmoid logistic model to estimate the time required for HBV DNA to fall below 5.3 $\log_{10}$ IU/mL. The present study addresses the same broad clinical problem using a different statistical framework, with an emphasis on patient-level heterogeneity and posterior uncertainty.

Bayesian Data Analysis provides a general framework for specifying and fitting Bayesian models [11]. Gelman also discussed prior distributions for variance parameters in hierarchical models, a relevant consideration when representing between-patient variability [12]. A hierarchical Bayesian approach is well suited to these data because each woman contributes repeated measurements, while the magnitude of the response can differ from one woman to another. The hierarchical structure lets the model borrow information across patients without forcing their trajectories to be identical.

Our aim was to describe longitudinal HBV DNA decline at the patient level, compare the two treatment groups within the fitted model, and examine how the decline changes between the early and later parts of follow-up. Additionally, we calculated the model-based threshold-crossing times and tested how sensitive the main treatment contrast was to the specified low-end observations.

## 2. Data and study design

The analysis used longitudinal HBV DNA measurements from 328 pregnant women with chronic HBV infection who received antiviral prophylaxis. Ninety-five women received telbivudine (LdT) and 233 received tenofovir disoproxil fumarate (TDF). The cleaned longitudinal dataset contained 1708 HBV DNA observations: 555 from the LdT group and 1153 from the TDF group.

The treatment comparison was observational rather than randomized. Time was centered at the recorded treatment-initiation week, so relative week 0 represents treatment initiation. The relative time ranged from 24 weeks before treatment initiation to 33 weeks afterward.

### 2.1. Published baseline characteristics

The published source study provides group-level baseline characteristics for the same 328 women [10]. These characteristics are summarized in Table 1. They show that the groups were similar in age and initial HBV DNA, but differed in several characteristics, including the timing of the first HBV DNA assessment and treatment initiation.

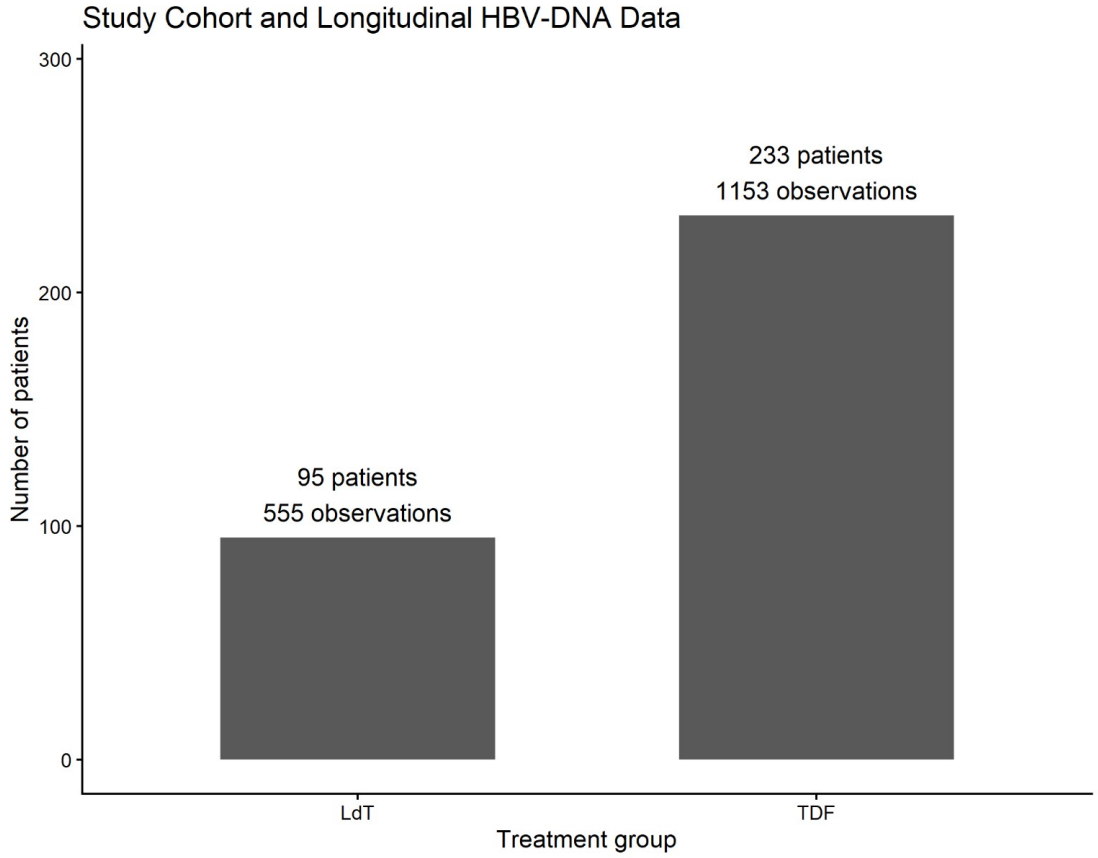



**Figure 1.** Overview of the longitudinal HBV DNA dataset and treatment groups. The analysis included 328 pregnant women, with 95 receiving LdT and 233 receiving TDF.

**Table 1.** Published baseline characteristics of the 328 women included in the source study.

| Characteristic | LdT ($n$ = 95) | TDF ($n$ = 233) | Total ($N$ = 328) | $p$ value |
|---|---|---|---|---|
| Age, median (range), years | 28 (19–38) | 28 (22–44) | 28 (19–44) | 0.551 |
| Childbirth gestational weeks, median (range) | 40 (37–43) | 39 (35–43) | 39 (35–43) | 0.020 |
| First parity, $n$ (%) | 83 (87.4) | 173 (74.2) | 256 (78.0) | 0.015 |
| Family history of chronic HBV, $n$ (%) | 39 (41.1) | 107 (45.9) | 146 (44.5) | 0.002 |
| Previous antiviral therapy, $n$ (%) | 14 (14.7) | 36 (15.5) | 50 (15.2) | 0.651 |
| HBeAg-positive chronic infection, $n$ (%) | 62 (65.3) | 155 (66.5) | 217 (66.2) | 0.603 |
| First HBV DNA assessment, median (range), weeks | 28 (8–32) | 17 (3–30) | 22 (3–32) | <0.001 |
| Treatment initiation week, median (range) | 28 (8–32) | 27 (14–32) | 26 (8–32) | 0.003 |
| ALT, median (range), U/L | 23 (6–75) | 24 (6–80) | 24 (6–80) | 0.415 |
| Initial HBV DNA, median (range), $\log_{10}$ IU/mL | 7.70 (5.55–9.18) | 8.05 (5.37–9.44) | 7.96 (5.37–9.44) | 0.241 |

*Note.* These are published group-level characteristics from Shen et al. [10]. They were not available as linked patient-level covariates in the longitudinal dataset used for the present analysis.

Although the published baseline table contains age, hepatitis B e-antigen (HBeAg) status, alanine transaminase (ALT), parity, family history, and previous antiviral treatment, these variables were not available for linkage to individual longitudinal trajectories. Therefore, we did not impute them or assign group-level values to individual patients. Instead, the adjusted analysis used the variables available at the patient level: baseline HBV DNA and treatment initiation week.

For the adjusted analysis, baseline HBV DNA was defined as the last HBV DNA measurement at or before treatment initiation. A qualifying baseline measurement was available for 94 of 95 LdT patients and 225 of 233 TDF patients. Thus, nine patients were excluded from the covariate-adjusted analysis, leaving 319 patients and 1671 observations.

## 3. Longitudinal HBV DNA dynamics

Individual trajectories showed substantial variation in both baseline viral load and the amount and timing of decline after treatment began. This variation was one of the main reasons for using a hierarchical patient-level model.

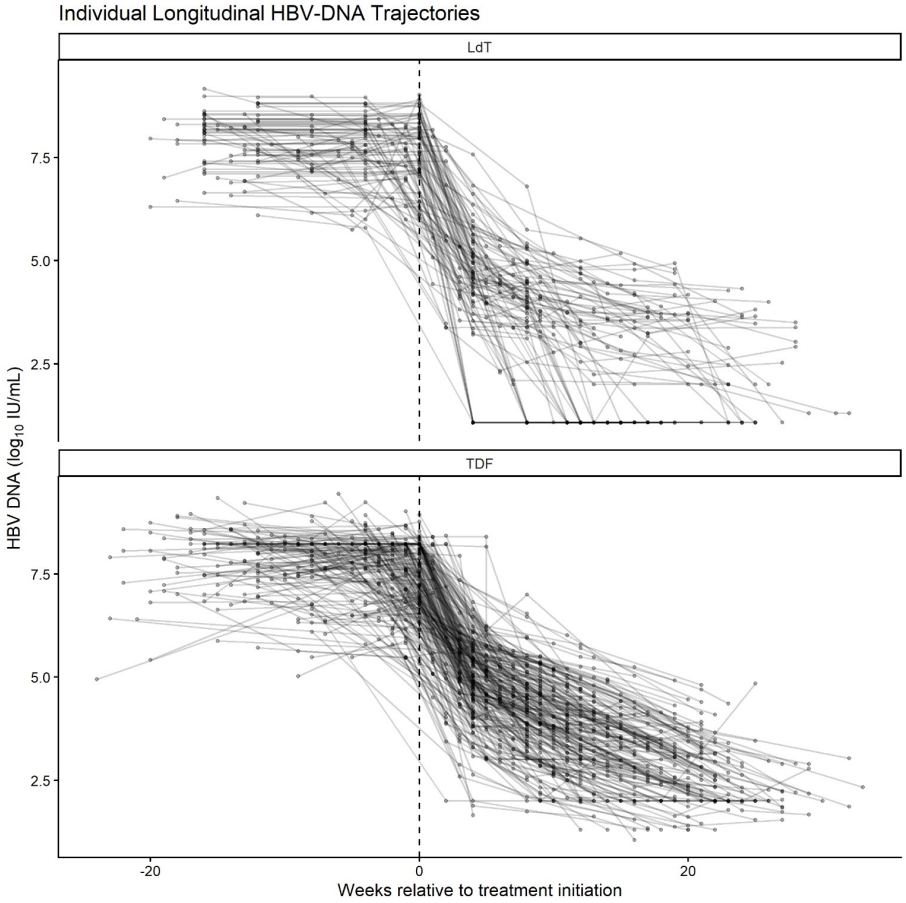



**Figure 2.** Individual longitudinal HBV DNA trajectories for women receiving LdT and TDF. Time is expressed relative to the treatment initiation, with week 0 representing the treatment-start week.

Figure 2 shows that the decline was generally rapid after treatment initiation and became slower at later follow-ups, although trajectories varied considerably from one patient to another.

## 4. Statistical model

### *4.1. Single-phase benchmark model*

As a benchmark, we first considered a single-phase exponential model. For patient $j$ and observation $i$, with time $t_{ij}$ measured relative to treatment initiation,

$$y_{ij} \sim \begin{cases} \mathcal{N}(c_j, \sigma_j), & t_{ij} \le 0, \\ \mathcal{N}\left(c_j \exp(-b_j t_{ij}), \sigma_j\right), & t_{ij} > 0. \end{cases} \tag{4.1}$$

Here, $c_j$ is a patient-specific fitted starting level, $b_j > 0$ is a decline-rate parameter, and $\sigma_j > 0$ is the residual standard deviation. Because the response variable is already on the $\log_{10}$ HBV DNA scale, these parameters describe a phenomenological statistical trajectory rather than a mechanistic model of circulating virions.

### *4.2. Continuous two-phase model*

The single-phase model imposes one post-treatment rate over the entire follow-up period. Following the motivation to examine early and later dynamics, we instead used a continuous two-phase model with a change in rate at week 12. For patient $j$,

$$m_j(t) = \begin{cases} c_j, & t \le 0, \\ c_j \exp(-b_{1j} t), & 0 < t \le 12, \\ c_j \exp(-12 b_{1j}) \exp[-b_{2j}(t-12)], & t > 12. \end{cases} \tag{4.2}$$

The observed response was modeled as follows:

$$y_{ij} \sim \mathcal{N}\left(m_j(t_{ij}), \sigma_j\right). \tag{4.3}$$

The first-phase rate $b_{1j}$ describes the decline during the first 12 weeks, while $b_{2j}$ describes the decline after week 12. The construction is continuous at the change point. The 12-week boundary is a pragmatic modeling choice used to compare early and later decline dynamics; it is not assumed to represent a biologically established transition.

### *4.3. Hierarchical structure and priors*

Patient-specific parameters were allowed to vary around the treatment-group means. For treatment group $g \in \{\text{LdT}, \text{TDF}\}$,

$$\log c_j = \mu_{\log c, g[j]} + \tau_{\log c, g[j]} z_{c,j}, \tag{4.4}$$

$$\log b_{1j} = \mu_{\log b_1, g[j]} + \beta_{1B} B_j + \beta_{1S} S_j + \tau_{\log b_1, g[j]} z_{b_1,j}, \tag{4.5}$$

$$\log b_{2j} = \mu_{\log b_2, g[j]} + \beta_{2B} B_j + \beta_{2S} S_j + \tau_{\log b_2, g[j]} z_{b_2,j}, \tag{4.6}$$

$$\log \sigma_j = \mu_{\log \sigma, g[j]} + \tau_{\log \sigma, g[j]} z_{\sigma, j}, \tag{4.7}$$

where $B_j$ is standardized baseline HBV DNA, and $S_j$ is standardized treatment initiation week. The latent patient effects were assigned standard normal distributions.

The group-level priors used in the Stan implementation were as follows:

$$\mu_{\log c, g} \sim \mathcal{N}(2, 1), \tag{4.8}$$

$$\mu_{\log b_1, g},\ \mu_{\log b_2, g} \sim \mathcal{N}(-2, 1), \tag{4.9}$$

$$\mu_{\log \sigma, g} \sim \mathcal{N}(0, 1), \tag{4.10}$$

$$\tau_{\log c, g}, \tau_{\log b_1, g}, \tau_{\log b_2, g}, \tau_{\log \sigma, g} \sim \mathcal{N}(0, 0.7), \tag{4.11}$$

Because the scale parameters are constrained to be non-negative in Stan, these correspond to half-normal priors on the between-patient standard deviations. The covariate coefficients were assigned weakly informative priors,

$$\beta_{1B}, \beta_{1S}, \beta_{2B}, \beta_{2S} \sim \mathcal{N}(0, 0.5). \tag{4.12}$$

### *4.4. Bayesian computation*

The models were fitted in R using Stan. Four Markov chain Monte Carlo (MCMC) chains were run with 2000 iterations per chain, including 1000 warm-up iterations, giving 4000 post-warm-up draws. The sampler used `adapt_delta = 0.95`. Hoffman and Gelman introduced the No-U-Turn sampler, which adaptively sets trajectory lengths in Hamiltonian Monte Carlo [13]. Betancourt provided a conceptual introduction to Hamiltonian Monte Carlo and its statistical foundations [14]. These methods were used to sample from the posterior distributions in this analysis.

Convergence and sampling behavior were assessed using $\widehat{R}$, effective sample size, divergent transitions, maximum tree depth, and the energy Bayesian fraction of missing information (E-BFMI). The use of $\widehat{R}$ and effective sample size follows modern recommendations for assessing MCMC convergence [15].

### *4.5. Threshold-crossing times*

For a threshold $H$, the posterior threshold-crossing times were calculated from each posterior draw. If the threshold was reached during the first phase, then

$$t_H = \frac{\log(c/H)}{b_1}. \tag{4.13}$$

If the threshold was reached after week 12, then the continuous second-phase expression was used. The threshold of 5.3 $\log_{10}$ IU/mL was emphasized because it is the threshold used in the source clinical study [10] and corresponds to approximately 200,000 IU/mL.

## 5. Results

### *5.1. Early and late decline rates*

The covariate-adjusted analysis included 319 patients and 1671 observations: 94 LdT patients and 225 TDF patients.

The posterior mean early-phase decline rate was 0.0964 week$^{-1}$ for LdT (95% CrI: 0.0862–0.1073) and 0.0752 week$^{-1}$ for TDF (95% CrI: 0.0714–0.0791). The posterior mean difference, defined as LdT minus TDF, was 0.0212 week$^{-1}$ (95% CrI: 0.0100–0.0332), with posterior probability 1.000 that the LdT rate was larger.

After week 12, the posterior mean decline rates were 0.0084 week$^{-1}$ for LdT (95% CrI: 0.0033–0.0155) and 0.0143 week$^{-1}$ for TDF (95% CrI: 0.0092–0.0199). The corresponding difference was −0.0059 week$^{-1}$ (95% CrI: −0.0132–0.0021), with a posterior probability 0.073 that the LdT rate was larger.

For both treatment groups, the posterior probability that the early-phase rate exceeded the late-phase rate was 1.000. The posterior mean early-minus- late differences were 0.0879 week$^{-1}$ for LdT and 0.0608 week$^{-1}$ for TDF.

**Table 2.** Covariate-adjusted posterior estimates of HBV DNA decline rates.

| Phase | LdT mean (95% CrI) | TDF mean (95% CrI) | LdT–TDF difference (95% CrI) |
|---|---|---|---|
| 0–12 weeks | 0.0964 (0.0862, 0.1073) | 0.0752 (0.0714, 0.0791) | 0.0212 (0.0100, 0.0332) |
| >12 weeks | 0.0084 (0.0033, 0.0155) | 0.0143 (0.0092, 0.0199) | −0.0059 (−0.0132, 0.0021) |

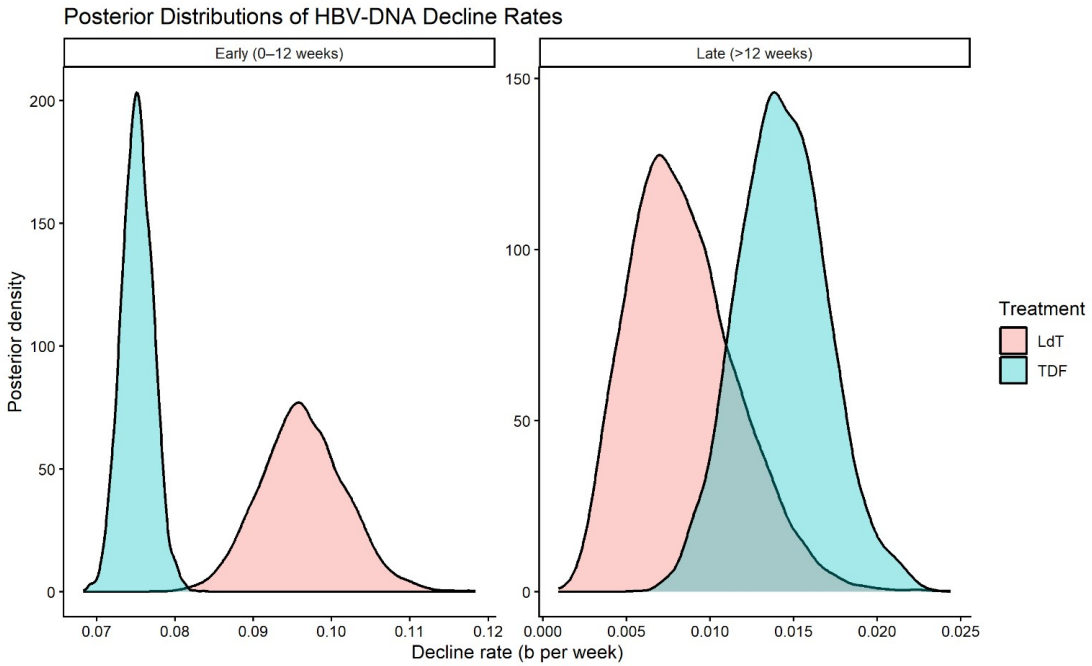



**Figure 3.** Posterior distributions of the early (0–12 weeks) and late (>12 weeks) HBV DNA decline rates for LdT and TDF. The distributions show the substantially faster early decline in both groups.

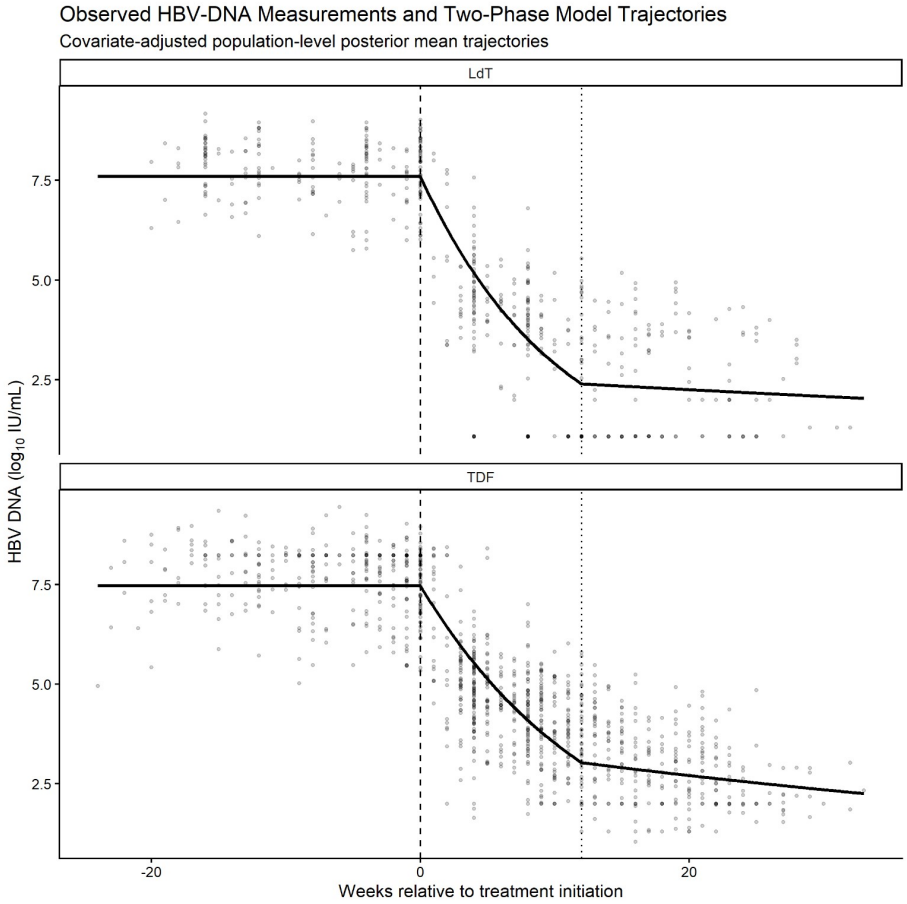



**Figure 4.** Observed HBV DNA measurements and population-level fitted trajectories from the covariate-adjusted continuous two-phase Bayesian model. The vertical reference lines indicate treatment initiation at week 0 and the week-12 transition.

The fitted trajectories in Figure 4 reproduce the broad pattern of a relatively rapid decline soon after treatment initiation, followed by a slower decline.

### *5.2. Associations with baseline HBV DNA and treatment timing*

In the adjusted model, the posterior mean coefficient for baseline HBV DNA in the early-rate model was −0.15 (95% CrI: −0.20 to −0.10), while the corresponding treatment-initiation-week coefficient was 0.05 (95% CrI: 0.00–0.09). For the late-rate model, the baseline HBV DNA coefficient was 0.47 (95% CrI: 0.15–0.83), while the treatment-initiation- week coefficient was −0.10 (95% CrI: −0.25–0.06).

These coefficients describe associations within the fitted model. Because the study was observational and important clinical covariates were unavailable, they should not be interpreted as causal effects.

### *5.3. Time to reach 5.3 $\log_{10}$ IU/mL*

In the covariate-adjusted two-phase model, the posterior mean estimated time to reach 5.3 $\log_{10}$ IU/mL was 3.75 weeks for LdT (95% CrI: 3.33–4.19) and 4.56 weeks for TDF (95% CrI: 4.30–4.84). The posterior probability that the threshold was reached earlier under LdT was 0.99875.

**Table 3.** Adjusted model-based threshold-crossing times.

| Threshold ($\log_{10}$ IU/mL) | LdT mean (weeks) | TDF mean (weeks) | $P(t_{LdT} < t_{TDF})$ |
|---|---|---|---|
| 5.3 | 3.75 | 4.56 | 0.99875 |
| 5.0 | 4.36 | 5.34 | 0.99950 |
| 4.0 | 6.68 | 8.31 | 0.99950 |
| 3.0 | 9.67 | 12.97 | 0.99975 |
| 2.0 | 36.43 | 41.96 | 0.73150 |

The 5.3 $\log_{10}$ IU/mL estimates are model-based population-level quantities rather than individual treatment recommendations. Additionally, the large uncertainty at the 2.0 threshold illustrates the difficulty of extrapolating far into the low viral-load range.

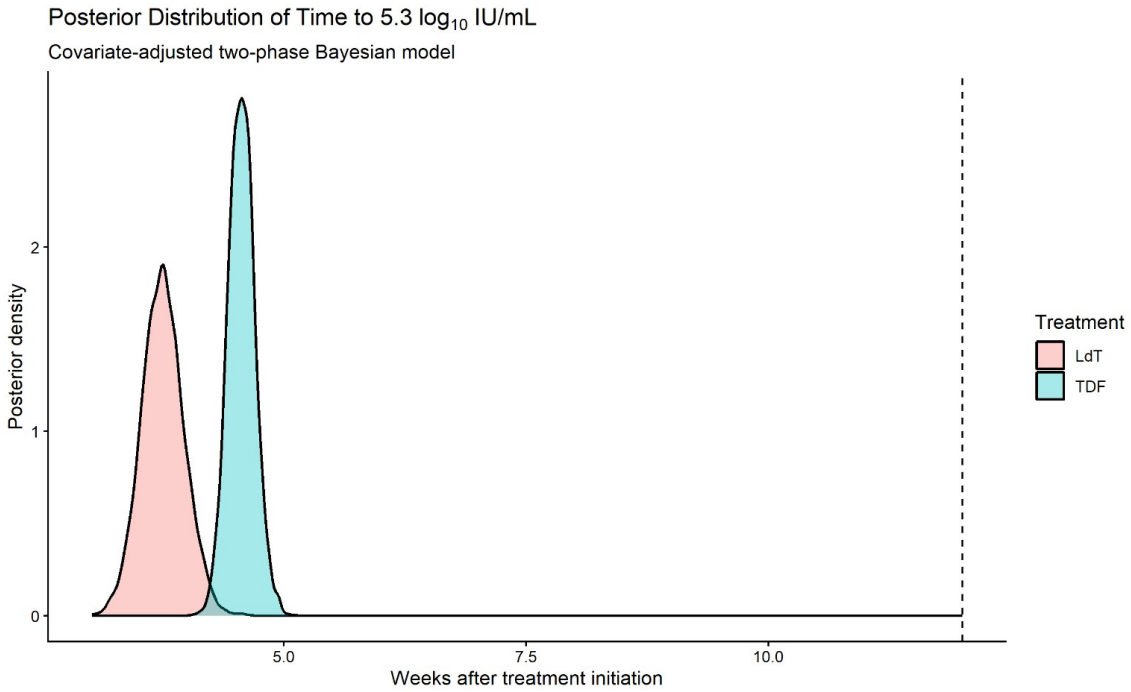



**Figure 5.** Posterior distributions of the estimated time to reach 5.3 $\log_{10}$ IU/mL under the covariate-adjusted two-phase model.

### *5.4. Posterior predictive assessment*

Posterior predictive replicated observations were generated from the **unadjusted** two-phase hierarchical model. Across 1708 observations used for this global predictive comparison, the correlation between the observed HBV DNA and the posterior predictive means was 0.9627 and the root mean square error (RMSE) was 0.6136 $\log_{10}$ units.

For comparison, the corresponding single-phase model gave a correlation of 0.9637 and an RMSE of 0.6174. Thus, the two-phase model reduced RMSE by only about 0.60% and did not improve the correlation. The two-phase formulation is therefore motivated mainly by the more informative description of changing decline dynamics rather than by a large gain in overall predictive accuracy.

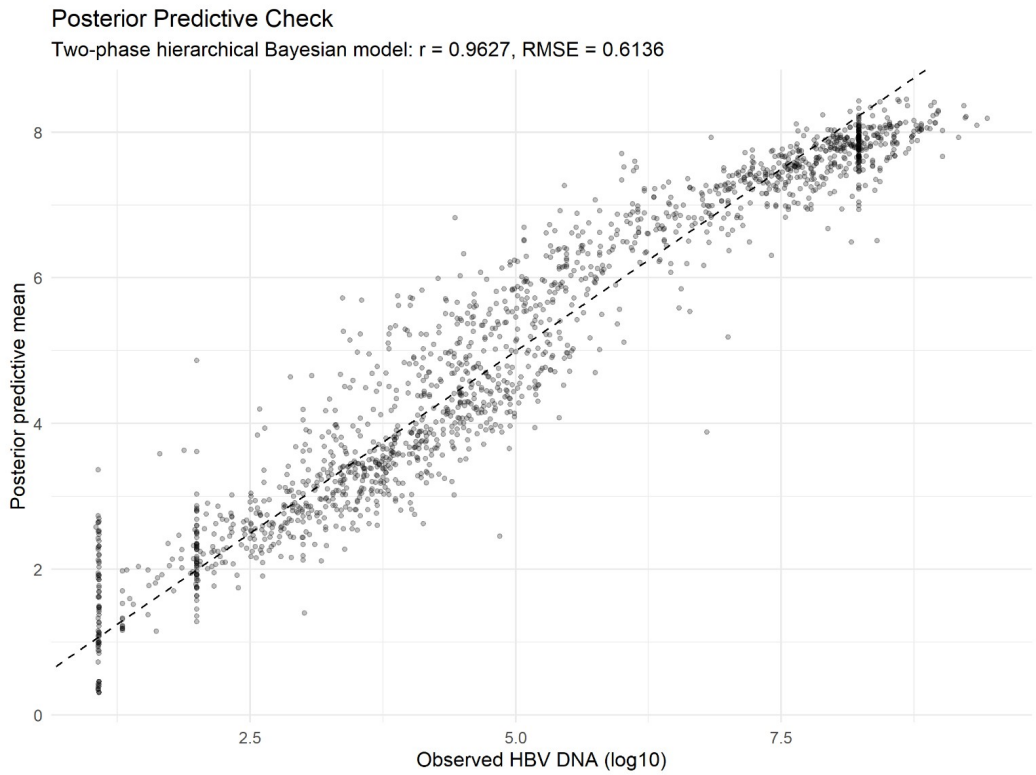



**Figure 6.** Posterior predictive check for the **unadjusted** two-phase hierarchical Bayesian model. Observed HBV DNA values are compared with posterior predictive means. The correlation was 0.9627 and the RMSE was 0.6136 $\log_{10}$ units.

### *5.5. Sensitivity analysis*

We repeated the adjusted analysis after excluding observations with HBV DNA values of 1.07 and 1.08 $\log_{10}$ IU/mL. These were treated as specified low-end observations for the sensitivity analysis; their exclusion does not by itself establish that they were censored by the assay.

After excluding these specified observations, the adjusted dataset contained 1592 observations from the same 319 patients. The posterior mean early-phase rate was 0.0799 week$^{-1}$ for LdT (95% CrI: 0.0731–0.0872) and 0.0752 week$^{-1}$ for TDF (95% CrI: 0.0713–0.0791). The early-phase difference was 0.0048 week$^{-1}$ (95% CrI: −0.0029–0.0131), with posterior probability 0.8825 that the LdT rate was larger.

For the late phase, the posterior mean rates were 0.0034 week$^{-1}$ for LdT (95% CrI: 0.0010–0.0071) and 0.0142 week$^{-1}$ for TDF (95% CrI: 0.0093–0.0198). The late-phase difference was −0.0108 week$^{-1}$ (95% CrI: −0.0170–−0.0049).

Therefore, the sensitivity analysis substantially weakened the early treatment group difference seen in the primary model. This result is an important qualification of the treatment comparison.

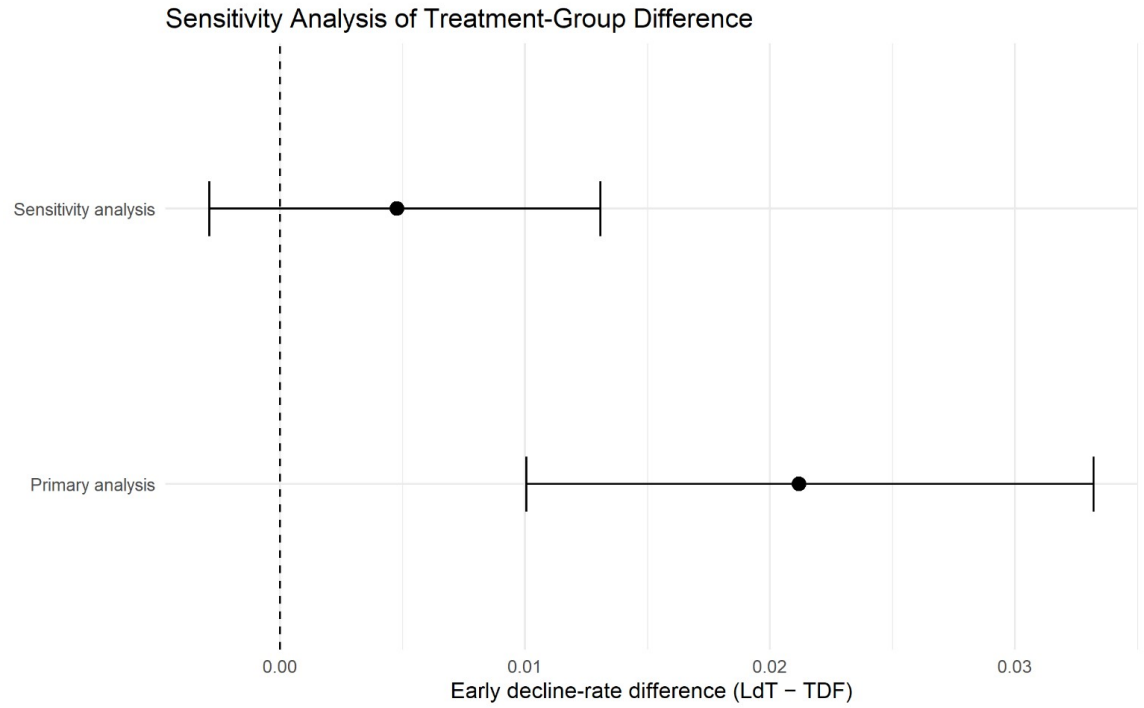



**Figure 7.** Sensitivity analysis of the early HBV DNA decline-rate difference between LdT and TDF. The primary estimate was 0.0212 week$^{-1}$ (95% CrI: 0.0100–0.0332), whereas the sensitivity estimate was 0.0048 week$^{-1}$ (95% CrI: −0.0029–0.0131).

### *5.6. Model diagnostics*

The primary covariate-adjusted two-phase model showed a generally stable sampling behavior. The maximum $\widehat{R}$ was approximately 1.017 and the minimum effective sample size was approximately 400. There were no divergent transitions, and the maximum tree depth was 6. E-BFMI values across the four chains were approximately 0.76, 0.81, 0.72, and 0.72.

## 6. Discussion

The clearest feature of the fitted trajectories was the change in decline rate over treatment time. In both groups, the HBV DNA fell relatively quickly during the first 12 weeks and then declined much more slowly. The hierarchical model also allowed the magnitude of this response to vary across patients.

The two-phase formulation was chosen because this pattern is visible in the longitudinal data. Tsiang et al. reported biphasic HBV clearance during adefovir therapy [6], and Wolters et al. examined viral dynamics during lamivudine therapy [7]. Perelson and Ribeiro reviewed HBV kinetics and mathematical modelling [8]. However, the 12-week boundary remains a pragmatic analysis choice rather than a claim that a biological transition occurs precisely at that time.

In the adjusted primary analysis, the early decline rate was higher for LdT than for TDF, with relatively high posterior certainty. However, the sensitivity analysis considerably changed the magnitude of the difference and produced a credible interval containing zero, meaning the apparent early treatment contrast is not fully robust to the handling of low-end HBV DNA observations.

Additionally, the results illustrate why treatment groups should not be summarized by a single overall ranking. In the primary model, the late-phase difference numerically favored TDF, but its credible interval included zero. In the sensitivity analysis, the late-phase contrast was more clearly shifted toward TDF. Taken together, these results do not support a claim that one treatment is generally superior to the other.

The results should not be read as a direct replication of Shen et al. [10]. That study used a sigmoid logistic model for the same clinical cohort, whereas we used patient-specific exponential trajectories within a hierarchical Bayesian model. Therefore, the numerical threshold times can differ because the two analyses describe the data in different ways. Instead, the purpose of the present model is to make patient heterogeneity, posterior uncertainty, and changes in decline rate explicit.

The 5.3 $\log_{10}$ IU/mL threshold times are model-based summaries of the fitted trajectories, not individual treatment recommendations. Prospective validation would be needed before using such estimates for clinical decision-making.

An important limitation is the observational nature of the treatment comparison. The published baseline table shows differences in treatment timing and several other characteristics, but individual-level information for age, HBeAg status, ALT, parity, family history, and previous antiviral therapy was not available for linkage to the longitudinal observations. Adjustment for baseline HBV DNA and the treatment initiation week can reduce some imbalances, but it cannot remove unmeasured or unavailable confounding. Therefore, the treatment-group contrasts should be regarded as model-based associations rather than causal effects.

Another limitation is the treatment of low-end measurements. The sensitivity analysis shows that these observations can have a noticeable influence on the estimated treatment contrast. Because the assay detection and quantification limits were not available in the patient-level data used here, we did not assume that the excluded observations were formally censored. Future work could model measurement limits explicitly if the assay characteristics were available.

Finally, the exponential model is phenomenological. It does not explicitly represent infected hepatocytes, intracellular viral processes, drug pharmacodynamics, or immune responses. More mechanistic or flexible models, combined with richer individual-level clinical data and external validation, could provide a useful next step.

## 7. Conclusions

We used a hierarchical Bayesian model to describe longitudinal HBV DNA dynamics during antiviral prophylaxis in pregnant women. Across both treatment groups, the fitted trajectories showed a pronounced early decline followed by a slower later phase.

The adjusted analysis estimated a higher early decline rate for LdT than for TDF, but this difference became much smaller when the specified low-end observations were removed. Together with the observational design and the limited availability of patient-level covariates, this sensitivity means that the analysis does not establish superiority of either treatment.

Therefore, the main value of the framework is descriptive: it provides a coherent way to represent patient heterogeneity, uncertainty, changing decline dynamics, and model-based threshold times while making the important limitations of the available data explicit.

## Ethics and data availability

This study is a secondary analysis of longitudinal data derived from the published cohort of Shen et al. [10]; no new participants were recruited for the present analysis. The source study reported institutional ethics approval (XJTU1AFCRC2020SJ-00125) and informed consent.

The individual-level longitudinal dataset used here was reconstructed from the published tables and supplementary materials of the source study. The authors were not provided with additional individual-level clinical covariates beyond those available in the published materials. Analysis code, model code, and derived data-processing materials can be made available by the authors subject to the conditions governing use of the original data.

## Reproducibility

The Bayesian analyses were performed in R and Stan using four MCMC chains, 2000 iterations per chain, 1000 warm-up iterations, and `adapt_delta = 0.95`. The analysis materials include the cleaned longitudinal dataset, Stan model code, R fitting scripts, and generated figures. The figures presented in this manuscript were generated from the corresponding R/RStan analyses and are included with the manuscript source materials.

## Use of AI tools declaration

The authors take full responsibility for the final content.

## Acknowledgments

The authors thank the University Research Facility in Big Data Analytics (UBDA) for computational resources, and the Smart Cities Research Institute and the Hong Kong Polytechnic University for financial support.

The work was partially supported by the Hong Kong Research Grants Council Theme-based Research Scheme fund (T24-508/22-N), the Hong Kong Research Grants Council General Research Fund (15307626), and the Hong Kong Research Grants Council Collaborative Research Fund (C5079-21G).

## Conflict of interest

The authors declare no conflicts of interest.

## References

1. European Association for the Study of the Liver, (2017) EASL 2017 Clinical Practice Guidelines on the management of hepatitis B virus infection. *J Hepatol* 67: 370–398. https://doi.org/10.1016/j.jhep.2017.03.021
2. World Health Organization, (2024) *Guidelines for the Prevention, Diagnosis, Care and Treatment for People with Chronic Hepatitis B Infection*, World Health Organization, Geneva, 2024.

3. Pan CQ, Duan Z, Dai E, Zhang S, Han G, Wang Y, et al. (2016) Tenofovir to prevent hepatitis B transmission in mothers with high viral load. *N Engl J Med* 374: 2324–2334. https://doi.org/10.1056/NEJMoa1508660
4. Brown Jr RS, McMahon BJ, Lok ASF, Wong JB, Ahmed AT, Mouchli MA, et al. (2016) Antiviral therapy in chronic hepatitis B viral infection during pregnancy: A systematic review and meta-analysis. *Hepatology* 63: 319–333. https://doi.org/10.1002/hep.28302
5. Nowak MA, Bonhoeffer S, Hill AM, Boehme R, Thomas HC, McDade H, (1996) Viral dynamics in hepatitis B virus infection. *Proc Natl Acad Sci* 93: 4398–4402. https://doi.org/10.1073/pnas.93.9.4398
6. Tsiang M, Rooney JF, Toole JJ, Gibbs CS, (1999) Biphasic clearance kinetics of hepatitis B virus from patients during adefovir dipivoxil therapy. *Hepatology* 29: 1863–1869. https://doi.org/10.1002/hep.510290626
7. Wolters LMM, Hansen BE, Niesters HGM, Zeuzem S, Schalm SW, de Man RA, (2002) Viral dynamics in chronic hepatitis B patients during lamivudine therapy. *Liver* 22: 121–126. https://doi.org/10.1034/j.1600-0676.2002.01579.x
8. Perelson, AS, Ribeiro RM (2004) Hepatitis B virus kinetics and mathematical modeling. *Semin Liver Dis* 24: 11–16. https://doi.org/10.1055/s-2004-828673
9. Ciupe SM (2018) Modeling the dynamics of hepatitis B infection, immunity, and drug therapy. *Immunol Rev* 285: 38–54. https://doi.org/10.1111/imr.12686
10. Shen M, He S, Yao N, Li R, Wang J, Zhong W, et al. (2025) Real-world clinical data-driven modelling on the initiation time of antiviral prophylaxis among pregnant women with chronic hepatitis B infection. *J Hepatol* 82: 816–825. https://doi.org/10.1016/j.jhep.2024.11.017
11. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB (2013), *Bayesian Data Analysis*, Chapman & Hall/CRC Press. https://doi.org/10.1201/b16018
12. Gelman A (2006) Prior distributions for variance parameters in hierarchical models. *Bayesian Anal* 1: 515–534. https://doi.org/10.1214/06-BA117A
13. Hoffman MD, Gelman A (2014) The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *J Mach Learn Res* 15: 1593–1623.
14. Betancourt M (2017) A conceptual introduction to Hamiltonian Monte Carlo, preprint, arXiv:1701.02434. https://doi.org/10.48550/arXiv.1701.02434
15. Vehtari A, Gelman A, Simpson D, Carpenter B, Bürkner PC, (2021) Rank-normalization, folding, and localization: An improved $\widehat{R}$ for assessing convergence of MCMC (with Discussion). *Bayesian Anal* 16: 667–718. https://doi.org/10.1214/20-BA1221




AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0)