

Retrospective Study

Time series analysis-based seasonal autoregressive fractionally integrated moving average to estimate hepatitis B and C epidemics in China

Yong-Bin Wang, Si-Yu Qing, Zi-Yue Liang, Chang Ma, Yi-Chun Bai, Chun-Jie Xu

Specialty type: Gastroenterology and hepatology**Provenance and peer review:**

Unsolicited article; Externally peer reviewed.

Peer-review model: Single blind**Peer-review report's scientific quality classification**Grade A (Excellent): A
Grade B (Very good): 0
Grade C (Good): C, C
Grade D (Fair): 0
Grade E (Poor): 0**P-Reviewer:** Huerta-Franco MR, Mexico; Kao JT, Taiwan; Vasily Isakov, Russia**Received:** August 5, 2023**Peer-review started:** August 5, 2023**First decision:** September 18, 2023**Revised:** September 28, 2023**Accepted:** October 23, 2023**Article in press:** October 23, 2023**Published online:** November 14, 2023**Yong-Bin Wang, Si-Yu Qing, Zi-Yue Liang, Chang Ma, Yi-Chun Bai**, Department of Epidemiology and Health Statistics, School of Public Health, Xinxiang Medical University, Xinxiang 453003, Henan Province, China**Chun-Jie Xu**, Beijing Key Laboratory of Antimicrobial Agents/Laboratory of Pharmacology, Institute of Medicinal Biotechnology, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100010, China**Corresponding author:** Yong-Bin Wang, MD, Researcher, Teacher, Department of Epidemiology and Health Statistics, School of Public Health, Xinxiang Medical University, No. 601 Jinsui Road, Hongqi District, Xinxiang 453003, Henan Province, China.
191035@xxmu.edu.cn

Abstract

BACKGROUND

Hepatitis B (HB) and hepatitis C (HC) place the largest burden in China, and a goal of eliminating them as a major public health threat by 2030 has been set. Making more informed and accurate forecasts of their spread is essential for developing effective strategies, heightening the requirement for early warning to deal with such a major public health threat.

AIM

To monitor HB and HC epidemics by the design of a paradigmatic seasonal autoregressive fractionally integrated moving average (SARFIMA) for projections into 2030, and to compare the effectiveness with the seasonal autoregressive integrated moving average (SARIMA).

METHODS

Monthly HB and HC incidence cases in China were obtained from January 2004 to June 2023. Descriptive analysis and the Hodrick-Prescott method were employed to identify trends and seasonality. Two periods (from January 2004 to June 2022 and from January 2004 to December 2015, respectively) were used as the training sets to develop both models, while the remaining periods served as the test sets to evaluate the forecasting accuracy.

RESULTS

There were incidents of 23400874 HB cases and 3590867 HC cases from January 2004 to June 2023. Overall, HB remained steady [average annual percentage change (AAPC) = 0.44, 95% confidence interval (95%CI): -0.94-1.84] while HC was increasing (AAPC = 8.91, 95%CI: 6.98-10.88), and both had a peak in March and a trough in February. In the 12-step-ahead HB forecast, the mean absolute deviation (15211.94), root mean square error (18762.94), mean absolute percentage error (0.17), mean error rate (0.15), and root mean square percentage error (0.25) under the best SARFIMA (3, 0, 0) (0, 0.449, 2)₁₂ were smaller than those under the best SARIMA (3, 0, 0) (0, 1, 2)₁₂ (16867.71, 20775.12, 0.19, 0.17, and 0.27, respectively). Similar results were also observed for the 90-step-ahead HB, 12-step-ahead HC, and 90-step-ahead HC forecasts. The predicted HB incidents totaled 9865400 (95%CI: 7508093-12222709) cases and HC totaled 1659485 (95%CI: 856681-2462290) cases during 2023-2030.

CONCLUSION

Under current interventions, China faces enormous challenges to eliminate HB and HC epidemics by 2030, and effective strategies must be reinforced. The integration of SARFIMA into public health for the management of HB and HC epidemics can potentially result in more informed and efficient interventions, surpassing the capabilities of SARIMA.

Key Words: Hepatitis; Seasonal autoregressive fractionally integrated moving average; Seasonal autoregressive integrated moving average; Prediction; Epidemic; Time series analysis

©The Author(s) 2023. Published by Baishideng Publishing Group Inc. All rights reserved.

Core Tip: This retrospective study used a seasonal autoregressive fractionally integrated moving average (SARFIMA) to monitor hepatitis B (HB) and hepatitis C (HC) epidemics, and its forecasting potential was then compared to that of the seasonal autoregressive integrated moving average (ARIMA) (SARIMA). The resulting forecast error rates under the SARFIMA were less than those under the SARIMA. The integration of SARFIMA into public health decision-making for the management of HB and HC epidemics can result in more informed interventions. The predicted HB totaled 9865400 [95% confidence interval (95%CI): 7508093-12222709] cases and HC totaled 1659485 (95%CI: 856681-2462290) cases in 2030, resulting in major challenges to eliminate hepatitis in China by 2030.

Citation: Wang YB, Qing SY, Liang ZY, Ma C, Bai YC, Xu CJ. Time series analysis-based seasonal autoregressive fractionally integrated moving average to estimate hepatitis B and C epidemics in China. *World J Gastroenterol* 2023; 29(42): 5716-5727

URL: <https://www.wjgnet.com/1007-9327/full/v29/i42/5716.htm>

DOI: <https://dx.doi.org/10.3748/wjg.v29.i42.5716>

INTRODUCTION

Hepatitis is a condition characterized by liver inflammation, which can be caused by various infectious viruses and noninfectious agents, resulting in a variety of health complications[1]. There are five primary strains of the hepatitis virus, namely, types A, B, C, D, and E. Hepatitis B (HB) and hepatitis C (HC), in particular, can progress into chronic diseases affecting millions of individuals, and they are the leading causes of liver cirrhosis, liver cancer, and mortality related to viral hepatitis[1,2]. Although significant achievements have been made in the prevention of HB and HC through the implementation of various strategies, these infections still pose significant health risks globally, affecting millions of people and causing a substantial burden on healthcare systems[3,4]. According to the World Health Organization (WHO), an estimated 296 and 58 million people were living with HB and HC, respectively, in 2019, and HB- and HC-related complications cause 820000 and 290000, respectively, deaths each year[1]. These numbers indicate the scale of the issue and emphasize the need for effective prevention and control strategies. Also, the WHO has set a goal of ending HB and HC as major public health threats by 2030 (compared with the 2015 baseline, new infections decreased by 30% by 2020 and 90% by 2030)[2]. China bears the greatest burden of HB virus (HBV) and HC virus (HCV) infections globally and is expected to play a pivotal role in achieving the global goal of eliminating HB and HC by 2030[4]. Accurate prediction aids in anticipating prospective scenarios and making proactive judgments, allowing policymakers to make educated decisions, develop strategies, and prepare for potential challenges and opportunities[5]. It has served as a valuable tool for mitigating risks, forming policies, and optimizing outcomes in different domains.

Time series analysis plays a crucial role in formulating hypotheses to comprehend the epidemics of different diseases and estimating the dynamics of observed phenomena, ultimately contributing to the establishment of a high-quality control system[5]. Various statistical methods, including generalized regression neural network[6], back propagation neural network[6], grey prediction[7], and Elman neural network[8], have been applied to predict the epidemics of infectious diseases. However, disease epidemics are frequently influenced by different factors, resulting in a combination of trends, seasonality, and randomness in the spread. Accordingly, the aforementioned statistical techniques may not adequately capture the complex and dynamic nature of epidemics, and it is challenging to generalize the models above

[5].

The seasonal autoregressive integrated moving average (SARIMA) has been the most commonly used model in different domains thanks to its simplicity, fast applicability, and ability to explain data[5,7,8]. Previous studies have also demonstrated the successful application of the SARIMA in estimating the prevalence, morbidity, and mortality of hepatitis[7], coronavirus disease 2019 (COVID-19)[5], hemorrhagic fever with renal syndrome (HFRS)[9], hand-foot-and-mouth disease (HFMD)[10], and tuberculosis[11]. The SARIMA can effectively capture the dynamic dependency structure by considering secular trends, periodic fluctuations, and random variations within a time series simultaneously[5]. Despite the attractive attributes of SARIMA, it is a representative of modelling short fluctuations in a series instead of the long term[12]. Additionally, using an integer difference in SARIMA for a series exhibiting long memory can cause over-differencing and the removal of valuable features, which harms forecasts[13]. In contrast, the seasonal autoregressive fractionally integrated moving average (SARFIMA) considers not only short memory, but also long memory, and such benefits as irregular fluctuations and complex seasonality by use of fractional differencing[12,14,15]. Also, the user-friendly nature of SARFIMA is relatively easy to explain to end-users and does not involve advanced mathematics or statistics. This enhances understanding and enables users to rely on the model for decision-making. The epidemiology of HB and HC is impacted by various factors, contributing to the epidemic patterns of long memory and complex dynamics [16]. Notwithstanding the advantages of SARFIMA, there is still a lack of exploration into how SARFIMA contributes to estimating HB and HC epidemics. Therefore, this study aimed: (1) To evaluate the usefulness of SARFIMA in monitoring HB and HC epidemics (projection into 2030) in mainland China; and (2) to assess the forecasting potential of SARFIMA compared to SARIMA.

MATERIALS AND METHODS

Data collection

The monthly incidence cases of HB and HC from January 2004 to June 2023 were provided by the Chinese Center for Disease Control and Prevention (<https://www.chinacdc.cn/>). The population numbers were obtained from the National Statistical Yearbook. All HB and HC cases were confirmed according to the HB and HC diagnosis criteria issued by The National Health Commission of the People's Republic of China (<http://www.nhc.gov.cn/wjw/s9491/wsbz.shtml>). HB and HC are notifiable diseases in China, and the confirmed cases required to be reported within 24 h. Any duplicate records were removed at the end of the same month. The SARIMA and SARFIMA were constructed using the data from January 2004 to June 2022, while the rest was the test set to indicate the predictive ability of both models. Because we required projection into December 2030 from July 2023 (90 data), the models were developed using the data between January 2004 and December 2015 to project the trends between January 2016 and June 2023 (90 data) to confirm the forecasting reliability.

SARIMA development

ARIMA is a widely used method that incorporates three components to model a time series: Autoregressive, differencing, and moving average models[17]. SARIMA is an extension of the ARIMA that includes a seasonal component, encompassing all the components of ARIMA and adding seasonal autoregressive, seasonal differencing, and seasonal moving average[18]. The model structure is represented as SARIMA (p, d, q) (P, D, Q)_s, where (p, d, q) represents the orders of the autoregressive, differencing, and moving average, respectively, (P, D, Q) represents the seasonal counterparts, and S is the seasonality. Selecting the appropriate values for $p, d, q, P, D,$ and Q in SARIMA involves four procedures. First, HB and HC data must be stationary. If the data is found to be non-stationary through an Augmented Dickey-Fuller (ADF) test, differencing should be applied to obtain stationary series[19]. Second, the initial orders for $p, q, P,$ and Q were identified by doing autocorrelation function (ACF) and partial ACF (PACF) analyses[17]. A series of attempts were then employed to identify the optimal combination by comparing Akaike's information criterion (AIC), corrected AIC (CAIC), Bayesian information criterion (BIC), and log-likelihood (LL)[18]. The one with lower values for AIC, BIC, and CAIC, along with higher LL was considered the preferred. Third, the errors from the best-fitting SARIMA were analyzed to verify if they behaved like a white noise series, which was done by the Ljung-Box Q test, ACF, and PACF[7]. Finally, once the model passed all the required statistical diagnoses, it could project into future periods.

SARFIMA development

Earlier values often significantly influence subsequent ones present in time series, demonstrating a concept known as hyperbolic decay (HD) time series[14]. By understanding the underlying mechanisms and developing accurate models for an HD series, researchers can extract important insights into complex systems and make data-driven predictions[13,14]. The SARIMA assumes an exponential decay of ACF, while the SARFIMA considers an HD pattern[9,13], indicating the presence of long-term memory. For this reason, the SARFIMA, with its added fractional differencing (df and Df), has been paid to considerable attention owing to its potential to capture persistent data[13,14]. df or Df ranges from -1 to 1 [15]: (1) A value < 0.5 shows a stationary series; (2) a value between -1 and -0.5 means an invertible series; (3) a value between -0.5 and 0 shows an anti-persistent series; (4) a value of 0 shows short memory and mean-reverting process in the series; (5) a value between 0 and 0.5 illustrates long-term positive dependence in the series; (6) a value between 0.5 and 1 suggests mean reverting but not stationarity in the series; and (7) a value of 1 signifies a unit root process[15]. The model structure is denoted as SARFIMA (p, d^1, q) (P, D^1, Q)_s, where $d^1 = d + df$ and $D^1 = D + Df$, whereby df or Df lies within $(-1, 0.5)$ as the fractional term and d or $D \geq 0$ is the integer term. The Hurst exponent (H), a measure of long-range dependency, often denotes the fractional differencing as df or $Df = H - 0.5$ [15]. The value of H falls between 0 and 1 , with $H < 0.5$ showing an

intermediate-memory series, $H = 0.5$ showing an uncorrelated series, and $H > 0.5$ indicating a long-memory series[15]. The H was computed using the rescaled range (R/S) method in this study[15]. As Veenstra's finding[14], the SARFIMA uses multiple starting values during its initial modelling phase. Consequently, it results in multiple modes, and the selection of the optimal one becomes a key step. From all possible modes, the one with the highest LL, alongside the lowest AIC and BIC, was the most considered[14]. The remaining steps of constructing the SARFIMA, such as parameter estimation and model diagnostics, were executed as described in SARIMA.

Statistical analysis

The Hodrick-Prescott method was employed to decompose the series into its trend and seasonality components[20]. The seasonal factor (SF) denotes the extent to which the incidence for a specific period deviates from the mean level ($SR > 1$ indicating a high-risk season; otherwise, a low-risk)[21], which was computed by multiplicative decomposition. The annual percentage change (APC) and average APC (AAPC) with a 95% confidence interval (95%CI) were calculated using Joinpoint (Version 4.8.0.1) to assess the changing trends[22]. The SARIMA and SARFIMA were implemented using the "forecast" and "arfima" packages in R (version 4.2.0, R Development CoreTeam, Vienna, Austria). Additionally, considering the significant impact of the COVID-19 pandemic on disease epidemics, a variable was thus created that was assigned a value of "1" between January 2020 and March 2023 and "0" between January 2004 and December 2019 to control for its impact on forecasting ability.

Evaluating the effectiveness of both models involved comparing the predicted values with the actual values of the test data using various performance metrics such as the mean absolute deviation (MAD), root mean square error (RMSE), mean absolute percentage error (MAPE), mean error rate (MER), and root mean square percentage error (RMSPE)[7]. A lower value on these metrics shows a more accurate forecast.

RESULTS

Data description

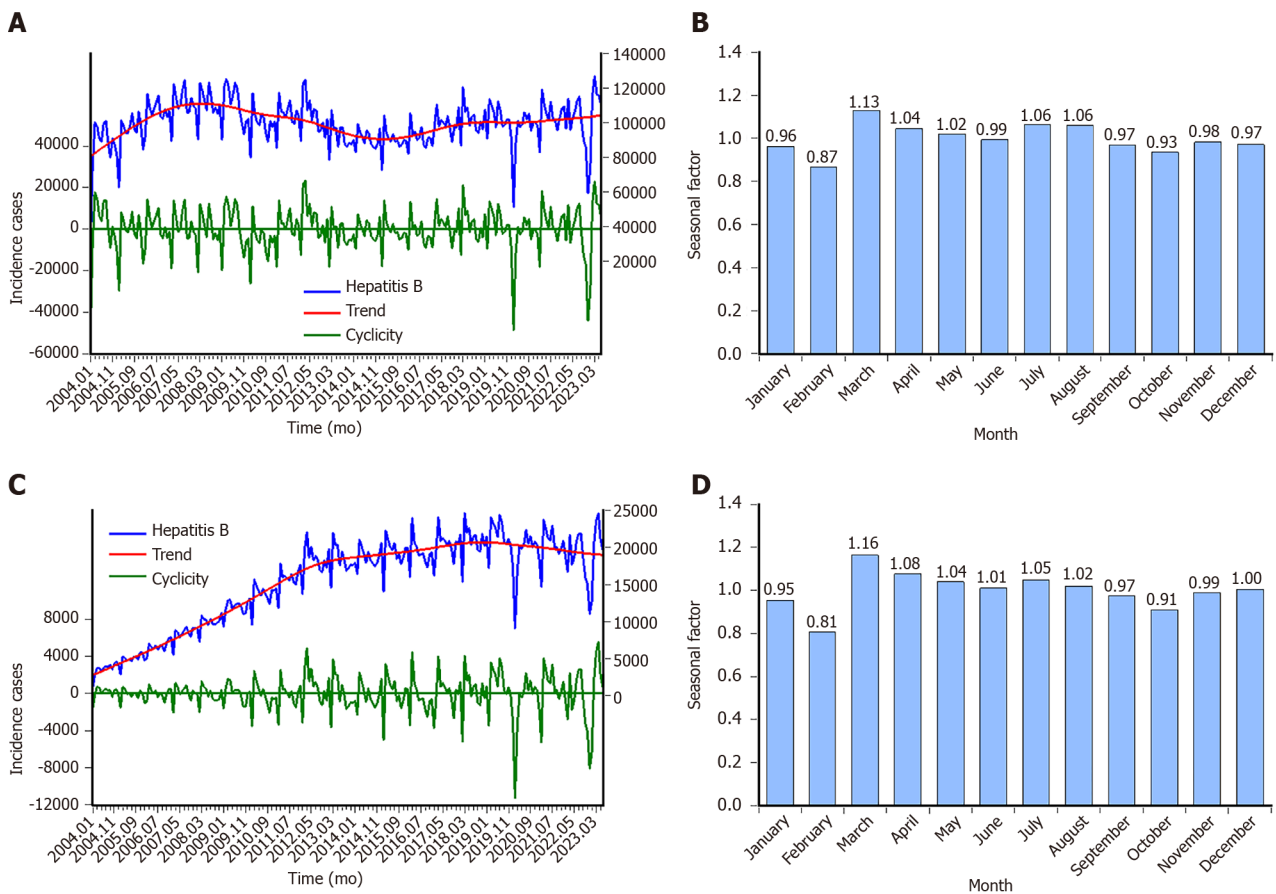
Between January 2004 and June 2023, China reported 23400874 HB cases (with incidence rates of 87.92 per 100000 persons per year and 7.33 per 100000 persons per month) and 3590867 HC cases (with incidence rates of 13.41 per 100000 persons per year and 1.12 per 100000 persons per month). **Figure 1** illustrates that HB exhibited an overall trend of stabilization (AAPC = 0.44, 95%CI: -0.94-1.84) but we observed a rise during 2004-2007 (APC = 9.13, 95%CI: 2.02-16.74), a decline during 2007-2014 (APC = -3.53, 95%CI: 2.02-16.74), and a steady level during 2014-2022 (APC = 0.86, 95%CI: -0.61-2.35) (**Figure 1A**). HC showed an overall trend of escalation (AAPC = 8.91, 95%CI: 6.98-10.88) but we noted an unpredictable upsurge during 2004-2007 (APC = 32.05, 95%CI: 20.31-44.94), a considerable increase during 2007-2012 (APC = 15.48, 95%CI: 11.34-19.78), a slight upturn during 2012-2019 (APC = 1.65, 95%CI: 0.08-3.24), and a reduction during 2019-2022 (APC = -4.30, 95%CI: -8.7-0.32) (**Figure 1C**). Based on **Figure 1B** and **D**, it is evident that HBV and HCV infections could occur all the year round. However, there were notable fluctuations in infection rates. March showed the highest infection rates, with an SF of 1.13 for HB and 1.16 for HC, signifying a peak. In contrast, February recorded the lowest infection rates, with SF values of 0.87 for HB and 0.81 for HC, indicating a trough. The infection rates for the other months remained relatively stable.

The preferred SARIMA

A stationary test for the HB incidence series between January 2004 and June 2022 generated an ADF = -0.29 ($P = 0.52$), and thus a non-stationary series was demonstrated. Given seasonality in the series, it was then seasonally differenced once (ADF = -6.12, $P < 0.001$), showing its stationarity. By observing the ACF and PACF plots of the differenced series (**Supplementary Figure 1**), a series of attempts were done, and we finally identified eight possible models with significant parameters (**Supplementary Table 1**). To ensure the selection of the optimal model, the "auto.arima" function in R was also run, ultimately leading to the automatic choice of SARIMA (1, 0, 2) (2, 0, 0)₁₂. By comparing the resulting nine models, it could be seen that SARIMA (3, 0, 0) (0, 1, 2)₁₂ was preferred as it gave the least values of AIC (4363.97), CAIC (4364.53), and BIC (4387.4), alongside the greatest value of LL (-2174.99). For the diagnoses of residuals, no correlations touched the significant bounds besides the ones at delays of 11 and 22 in the ACF and PACF plots (**Supplementary Figure 2A**) and $P = 0.49$ for the Ljung-Box Q statistic indicated an unrelated series of residuals, meaning that this preferred model sufficiently fits the data, which could be used to project into the next 12 data (**Figure 2A**). Likewise, following the modelling methods, SARIMA (1, 0, 2) (2, 1, 0)₁₂ (AIC = 2748.41, CAIC = 2749.08, BIC = 2765.71, and LL = -1368.21), SARIMA (3, 0, 0) (0, 1, 1)₁₂ (AIC = 3636.75, CAIC = 3637.17, BIC = 3656.84, and LL = -1812.38), and SARIMA (3, 1, 0) (2, 1, 0)₁₂ (AIC = 2022.09, CAIC = 2022.77, BIC = 2219.34, and LL = -1095.05) were determined as the optimal specifications for the 90-step-ahead forecasts of HB (**Figure 2B**), 12-step-ahead forecasts of HC (**Figure 3A**), and 90-step-ahead forecasts of HC (**Figure 3B**), respectively (statistical checks for the forecast errors are illustrated in **Supplementary Figure 2B-D**).

The preferred SARFIMA

The resulting R/S = 0.73 for the HB incidence series between January 2004 and June 2022 indicated that the series exhibited long-range dependence, and thus it was well suited to build SARFIMA. Subsequently, based on the fitting steps, the SARFIMA (3, 0, 0) (0, 0.449, 2)₁₂ with one mode tended to be indicated to be the preferred because it yielded the least values of AIC (4014.66) and BIC (4048.69), along with the maximum value of LL (-1997.33). **Supplementary**



DOI: 10.3748/wjg.v29.i42.5716 Copyright ©The Author(s) 2023.

Figure 1 Time series graphs displaying the incidence cases, trend, cyclicity, and seasonality of hepatitis B and C. A: Hepatitis B incidence cases and the decomposed trend and cyclicity by Hodrick-Prescott (HP) method; B: Seasonal factor for hepatitis B series; C: Hepatitis C incidence cases and the decomposed trend and cyclicity by HP method; D: Seasonal factor for hepatitis C series. It is clear that hepatitis B and C have the same seasonal patterns, a peak in March and a trough in February, and the other months remain relatively stable.

Figure 3A presents the ACF and PACF analyses for the residuals, showing that all spikes were within the significance bounds and $P = 0.48$ for the Ljung-Box Q statistic pinpointed that the forecast errors belonged to a white noise series. Accordingly, this preferred model could make a forecast for the 12-holdout data (Figure 3A). In the same way, SARFIMA (1, 0, 2) (2, 0.454, 0)₁₂ (R/S = 0.76, AIC = 2610.36, BIC = 2637.09, and LL = -1296.18), SARFIMA (3, 0, 0) (0, 0.428, 1)₁₂ (R/S = 0.84, AIC = 3257.01, BIC = 3287.63, and LL = -1619.5), and SARFIMA (3, -0.155, 0) (2, -0.274, 0)₁₂ (R/S = 0.83, AIC = 2036.61, BIC = 2063.34, and LL = -1009.3) were identified as the best model for the 90-step-ahead forecasts of HB (Figure 2B), 12-step-ahead forecasts of HC (Figure 3A), and 90-step-ahead forecasts of HC (Figure 3B), respectively (the resulting modes and the required diagnoses for the forecast errors are given in Supplementary Tables 2-4 and Supplementary Figure 3B-D).

Assessing predictive accuracy and reliability

The data in Table 1 reveals the forecasting accuracy of both SARIMA and SARFIMA. Notably, the SARFIMA exhibited smaller values of MAD, MAPE, RMSE, MER, and RMSPE. Figure 2A and B and Figure 3A and B demonstrate a comparison between the forecasts generated by both models and the observed values. The SARFIMA depicted a closer resemblance to the actual trends and seasonality in comparison to the SARIMA. These results meant that the SARFIMA outperformed the SARIMA. Moreover, two sensitivity analyses were conducted to examine the influence of the age of the affected population and the cultural patterns during the spring season on the predictive quality of SARFIMA. These analyses indicated that the SARFIMA consistently produced lower forecasting error rates compared to the SARIMA (Supplementary Tables 5 and 6). This reinforces the robustness of the SARFIMA. Consequently, a projection into 2030 for HB and HC epidemics was done by identifying the optimal SARFIMA (1, 0, 1) (0, 0.438, 1)₁₂ and SARFIMA (1, 0.429, 1) (2, 0, 0)₁₂, respectively, on the whole data. The resulting results indicated that HB would reach a plateau in the upcoming years (Figure 2C), and the forecasts totaled 9865400 (95%CI: 7508093-12222709) incidents, with a yearly average of 1233175 (95%CI: 938512-1527839) incidents (Table 2); HC would begin to recede in the next years (Figure 3C), and the forecasts totaled 1659485 (95%CI: 856681-2462290) incidents, with an annualized average of 207436 (95%CI: 107085-307786) incidents (Table 2).

Table 1 Comparison of the forecasting ability under the seasonal autoregressive integrated moving average and the seasonal autoregressive fractionally integrated moving average

Metrics	Hepatitis B		Hepatitis C	
	SARIMA	SARFIMA	SARIMA	SARFIMA
12-step ahead forecasts				
MAD	16867.708	15211.939	3355.664	3245.308
MAPE	0.191	0.173	0.208	0.201
RMSE	20775.123	18762.935	4093.553	3940.966
MER	0.165	0.149	0.178	0.172
RMSPE	0.267	0.248	0.285	0.279
90-step ahead forecasts				
MAD	13423.963	10446.219	2246.348	1586.973
MAPE	0.134	0.106	0.133	0.096
RMSE	15289.986	12577.033	3379.293	2482.687
MER	0.134	0.105	0.112	0.079
RMSPE	0.155	0.137	0.241	0.192

SARIMA: Seasonal autoregressive integrated moving average; SARFIMA: Seasonal autoregressive fractionally integrated moving average; MAD: Mean absolute deviation; MAPE: Mean absolute percentage error; RMSE: Root mean square error; MER: Mean error rate; RMSPE: Root mean square percentage error.

Table 2 Projection into 2030 for hepatitis B and C incidence cases

Time (yr)	Hepatitis B		Hepatitis C	
	Forecasts	95%CI	Forecasts	95%CI
2023	1240031	1041413-1438650	210742	141933-279552
2024	1262930	994670-1531191	209337	149760-268915
2025	1245681	953797-1537566	213123	138049-288198
2026	1234951	929339-1540563	209276	117296-301256
2027	1227438	912670-1542207	208070	101135-315004
2028	1221949	900585-1543314	205350	84304-326395
2029	1217813	891418-1544208	203073	69221-336926
2030	1214607	884203-1545011	200514	54983-346044
Mean	1233175	938512-1527839	207436	107085-307786
Total	9865400	7508093-12222709	1659485	856681-2462290

95%CI: 95%confidence interval.

DISCUSSION

Hepatitis poses a major threat to public health globally[2,16]. Analyzing and predicting epidemics are of great importance for the prevention and control. This study represents an important contribution to the field as it is the first to explore the efficacy of SARFIMA in monitoring HB and HC epidemics, while also comparing its predictive accuracy with that of SARIMA. The results supported our initial hypothesis, demonstrating the SARFIMA as a more comprehensive approach for capturing the epidemic dynamics of HB and HC compared to the SARIMA. Importantly, the forecasting robustness was confirmed by our further sensitivity analyses (Supplementary Tables 5 and 6). Also, previous work indicated that the SARFIMA showcased a good performance in forecasting costs[13], road fatality rate[15], CO₂ emission[23], temperature [14], and stock markets[24]. These findings provide further validation for the effectiveness of SARFIMA as a promising alternative in monitoring the spread of HB and HC. Moreover, the application of SARFIMA can superbly contribute to

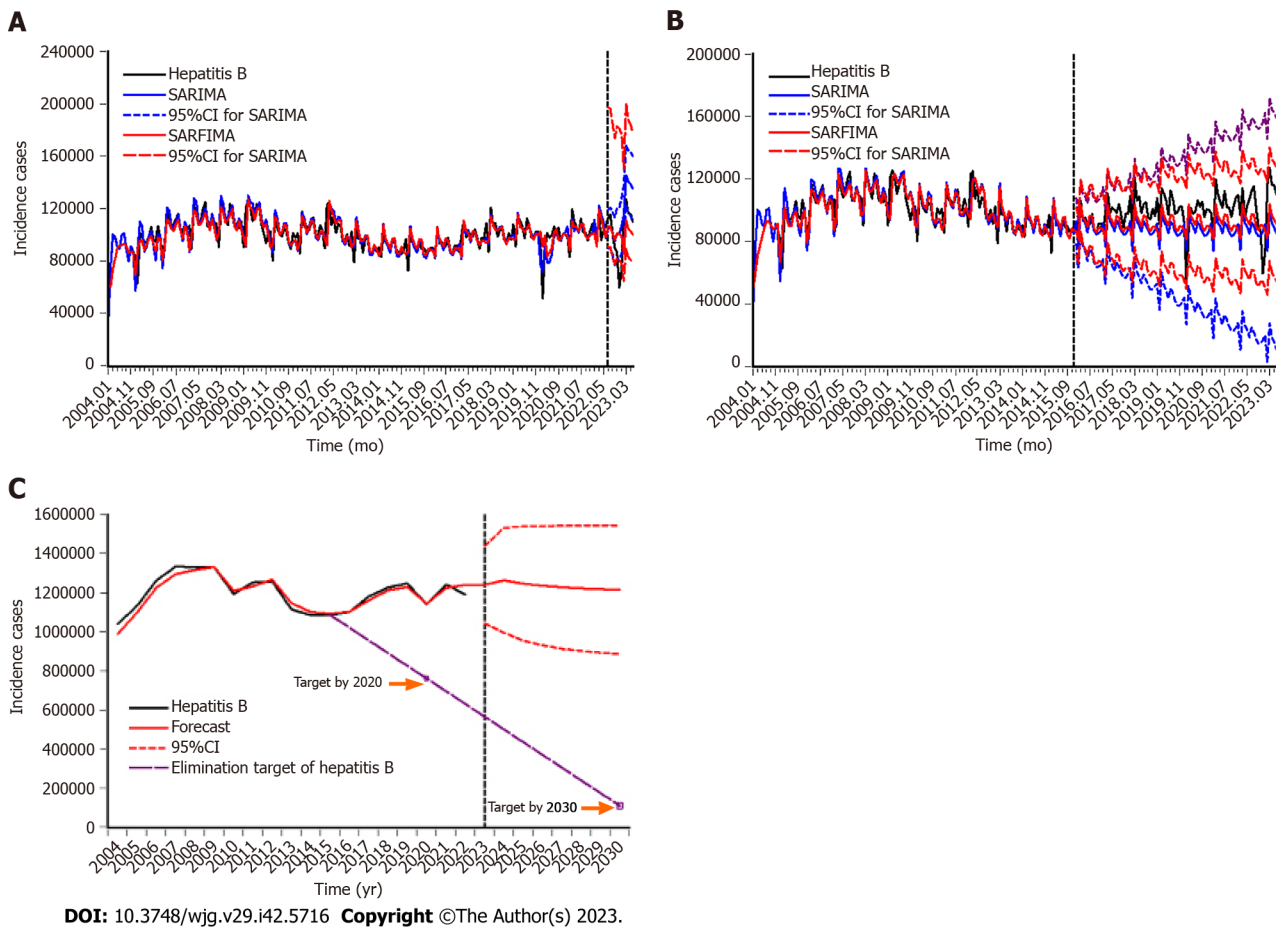


Figure 2 Comparison of the hepatitis B incidence cases with the forecasts under the seasonal autoregressive integrated moving average and seasonal autoregressive fractionally integrated moving average. A: Comparison of the 12-step ahead forecasts under the seasonal autoregressive integrated moving average (SARIMA) and seasonal autoregressive fractionally integrated moving average (SARFIMA); B: Comparison of the 90-step ahead forecasts under the SARIMA and SARFIMA; C: Projection into 2030 under the SARFIMA. The predictive results are plotted in the right of the virtual vertical line. It seemed that the forecasts under the SARFIMA are in closer proximity to the actual hepatitis B incidence values and the predictive hepatitis B incidence curves would reach a plateau in the upcoming years. 95%CI: 95% confidence interval; SARFIMA: Seasonal autoregressive fractionally integrated moving average; SARIMA: Seasonal autoregressive integrated moving average.

guiding the intensity and type of public health measures. For example, if the model clearly showed an upsurge in the midst of receding HB and HC epidemics, it would suggest the effectiveness of measures currently in place. Instead, if the SARFIMA predicted a decline despite increasing HB and HC epidemics, the need for further or optimized measures should be heightened. These practical and actionable insights hold great promise for SARFIMA in monitoring and controlling HB and HC epidemics.

The SARIMA, as a reliable method for analyzing and forecasting time series data, has gained significant recognition within the fields of economics, finance, meteorology, and healthcare[11,17,18]. It can adapt to different time series patterns by adjusting the orders of nonseasonal and seasonal terms that incorporate trend, seasonal, and random components into the modelling framework, and can often generate acceptable forecasts[5,7,8]. As evidenced by our study, the SARIMA generated a relatively accurate forecast despite its inferiority to the SARFIMA in terms of predictive quality. The SARIMA has been indicated to be satisfactory in predicting the spread of various infectious diseases such as hepatitis [7], COVID-19[5], HFRS[9], HFMD[10], and tuberculosis[11]. Also, analysts well-versed in the underlying principles and the procedural steps involved with the SARIMA have been able to generate informed and precise forecasts, thereby assisting decision-making and planning efforts in containing the spread of diseases. Notwithstanding this, the SARIMA has indicated to be effective in capturing regular short-run dynamics and simple seasonality, and it often generates over-differencing[13,18]. By contrast, the SARFIMA combining the strengths of SARIMA and fractional integration can capture long-range dependence, handle complex seasonal patterns, accommodate both stationarity and non-stationarity, provide reliable parameter estimates, and offer robustness and stability[14,15,23], enabling it an invaluable tool for monitoring HB and HC epidemics, contributing to more informed decision-making and improved understanding of complex temporal dynamics. Considering that time series analysis is a crucial aspect of forecasting that combines various factors and the comprehensive effects of uncertain variables into a time variable, which is cost-effective and widely applicable in practice [25], promoting the adoption of SARFIMA can contribute to the improved accuracy and reliability of modelling and forecasting other infectious diseases. However, it is crucial to further validate its generalization. It is also worth mentioning that recent studies have unveiled satisfactory applications of alternative models such as Bayesian structural

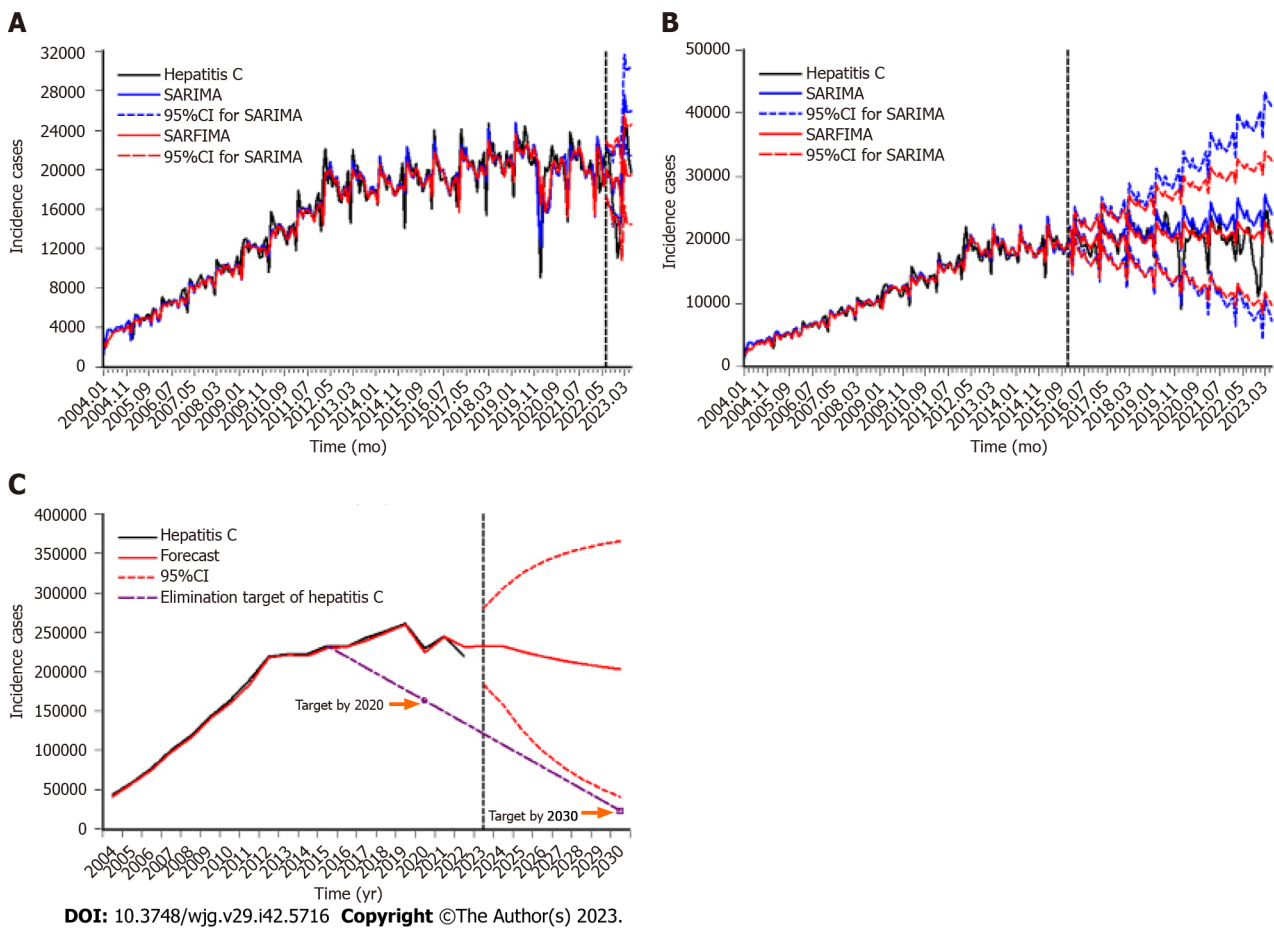


Figure 3 Comparison of the hepatitis C incidence cases with the forecasts under the seasonal autoregressive integrated moving average and seasonal autoregressive fractionally integrated moving average. A: Comparison of the 12-step ahead forecasts under the seasonal autoregressive integrated moving average (SARIMA) and seasonal autoregressive fractionally integrated moving average (SARFIMA); B: Comparison of the 90-step ahead forecasts under the SARIMA and SARFIMA; C: Projection into 2030 under the SARFIMA. The predictive results are plotted in the right of the virtual vertical line. It appeared that the forecasts under the SARFIMA are in closer proximity to the actual hepatitis C incidence values and the predictive hepatitis C incidence cases would begin to recede in the next years. 95%CI: 95% confidence interval; SARFIMA: Seasonal autoregressive fractionally integrated moving average; SARIMA: Seasonal autoregressive integrated moving average.

time series and innovation state-space framework for assessing the epidemics of diseases[26,27]. Accordingly, additional studies focus on comparing and confirming the forecasting performance of these models alongside the SARFIMA.

Different from the global declined trend in HB and HC incidences[2], an overall increase at an average rate of 0.44% for HB and 8.91% for HC per year was noted in our study, also consistent with earlier studies in Guangxi[8], China[4], and Pakistan[28]. There are several possible explanations for this trend. First, a study revealed that only around 20% of individuals infected with HBV in China were diagnosed[29]. While the gradual improvements in surveillance systems and improved diagnostic capabilities contribute to the detection of more cases[30]. Second, despite the successful implementation of the neonatal HB vaccination program, which has led to a noticeable decline in HB incidence and transmission, the prevalence of HB and HC remains alarmingly high in China. This is primarily due to the extensive population, approximately 296 and 58 million individuals, living with HB and HC, respectively. Of particular concern is the about 5% prevalence of HB surface antigen (HBsAg) and 2% prevalence of anti-HCV in women of reproductive age [31,32], with mother-to-child transmission being a significant risk. This poses a substantial challenge in terms of advancing diagnostic coverage, eradicating HBV and HCV infections, and reducing mortality. Third, despite the high curative rate achieved by direct-acting antiviral agents (DAAs) treatment for patients with HC[33], China's actual diagnostic rate for HC is only 2.1%[34]. Due to this low diagnostic rate, a smaller number of HC patients receive antiviral therapy, contributing to the ongoing spread of HC from high-risk populations to the general population. Lastly, coinfections of sexually transmitted diseases (STDs), HB, and HC are common and significantly contribute to long-term morbidity and mortality. In recent years, there has been a rapid increase in the incidence of STDs in China[35]. The high rates of HB and HC among individuals with STDs are predominantly driven by injection drug use (IDU) and sexual transmission, particularly among men who have sex with men (MSM)[36]. A significant decline in HB incidence during 2007-2014 can be attributed to the incorporation of HB prevention and treatment into the "Eleventh Five-Year" and "Twelfth Five-Year" plans in China. Comprehensive measures such as strengthening vaccination, enhancing public awareness and education, and conducting training programs were implemented[4,37]. Previous studies indicated that there was an upturn in HB cases in China since 2016/2017[7,8], aligning with our results, possibly attributable to the

accelerated urbanization process, a significant increase in the migrant population, a rapid rise in co-infections with HB, and the heavy economic burden. HC has insidious onset and nonspecific symptoms, making it difficult to be detected in the early stages. Although there is no effective vaccine for HC so far, with the gradual expansion of monitoring and testing coverage and the reporting of several outbreaks, more people have undergone screening, resulting in a rapid increase in confirmed cases during 2004-2012[30], and this matched well with our findings. Recent years have witnessed a decline in HC, which may be associated with the strict screening of blood donors and the standardized management of blood products, the comprehensive implementation of monitoring, early warning, intervention, and assessment measures for infectious diseases, the improvement in public knowledge-attitude-behavior regarding hepatitis, the continued optimization of policies, and the improvement of medical insurance[4,30]. Besides, according to the predicted figures for 2030, HB reached a plateau and HC receded. Thus, it can be said that the elimination of HB and HC by 2030 under current interventions faces enormous challenges. Therefore, addressing these challenges requires a comprehensive approach that encompasses prevention, diagnosis, and treatment[2,4]. First, efforts should be focused on expanding the reach of the neonatal HB vaccination program to ensure that all infants are protected from HBV infection[2,4]. Additionally, targeted vaccination campaigns should be implemented to reach vulnerable populations, such as women of reproductive age, in order to reduce the risk of mother-to-child transmission[3]. Second, it is imperative to improve the diagnostic rate of both HBV and HCV infections in China. This can be achieved through increased awareness and education among healthcare professionals, as well as the general population, regarding the importance of early detection and screening[2,4]. Furthermore, the establishment of screening programs in high-risk areas and the provision of affordable and accessible diagnostic tests can significantly contribute to improving the diagnostic rates. Third, it is crucial to strengthen the healthcare infrastructure and expand access to antiviral therapies for both HB and HC[2,4]. This can be achieved by training healthcare professionals in the latest treatment guidelines and protocols, as well as ensuring the availability and affordability of antiviral drugs. Lastly, efforts should be made to reduce the stigma associated with these diseases, as this can act as a barrier to seeking medical help and receiving appropriate treatment[2,4].

HB and HC show a seasonal profile in this study, with a trough in February and a peak in March, consistent with prior reports[7]. The seasonal trough during Spring Festival is largely attributed to people's reluctance to seek medical treatment and under-reporting that is more severe than in other months. However, the seasonal peak is associated with large-scaled population movement after Spring Festival and increased participation in various entertainment activities during the holiday. In China, although the primary mode of HB and HC transmission has shifted from drug use to sexual transmission, injecting drugs and engaging in unprotected high-risk sexual activities following drug use are common practices[38]. People's behavior is evolving to recognize the potential transmission risks of HCV/HBV, such as unprotected sexual contacts and IDU[38]. This shift is resulting in a growing demand for testing, as individuals are becoming more conscious of the need to check their status after festive events, leading to an increase in patients during this period. However, some studies have also indicated that HB and HC incidences follow the epidemiological characteristics of blood-borne and sexually transmitted diseases, with less pronounced seasonality[30].

Shortcomings also need to be considered in this study. First, data was taken from a passive monitoring system, and under-reporting is inevitable. Second, regional heterogeneity varied greatly in HB and HC incidences in China, and the region-specific forecasting ability of SARFIMA may require additional validation. Third, to capture timely information, the model requires integrating new data duly. Fourth, limited data may not exhibit long-term dependencies, and thus a series comprising 100 or more samples is recommended in application[17]. Fifth, due to the lack of available data pertaining to the nutritional status of the population, comorbidities such as diabetes and hypertension, the specifics of daily water consumption, and immunological resistance related to the genetic traits of the Chinese population, we are unable to provide a more detailed analysis of how these factors may impact the observed results. Lastly, as to whether the SARFIMA is transferable to monitor HB and HC in other study regions or other infectious diseases, verification is warranted.

CONCLUSION

Overall, HB remains steady while HC is rising in China, and both exhibit a seasonal pattern, a peak in March and a trough in February. Under current interventions, additional feasible and effective control strategies (e.g., expanding the scope of adult HB vaccination, a breakthrough in vaccination for HC, preventing mother-to-child transmission, investigating high-risk factors, implementing standardized antiviral treatment in rural areas, and enhancing health education and promotion) require to be designed to ensure the elimination of HB and HC by 2030. The SARFIMA provides a more sophisticated and adaptable framework for capturing intricate patterns and interdependencies in monitoring HB and HC epidemics, as opposed to SARIMA. This ultimately leads to enhanced forecasting capabilities and a deeper comprehension of the underlying process. Consequently, the integration of SARFIMA into public health decision-making for the management of HB and HC epidemics can result in more informed and efficacious interventions.

ARTICLE HIGHLIGHTS

Research background

Hepatitis B (HB) and hepatitis C (HC) have the largest burden in China, and a goal of eliminating them as a major public health threat by 2030 has been raised.

Research motivation

Accurate prediction helps to anticipate possible scenarios and make proactive choices, enabling policymakers to make informed decisions, plan strategies, and prepare for potential challenges and opportunities.

Research objectives

This study aimed to evaluate the usefulness of seasonal autoregressive fractionally integrated moving average (SARFIMA) in monitoring HB and HC epidemics (projection into 2030) in mainland China and to assess the forecasting potential of SARFIMA compared to seasonal autoregressive integrated moving average (SARIMA).

Research methods

The monthly incidence cases of HB and HC from January 2004 to June 2023 were obtained. Then, the two periods (from January 2004 to June 2022 and from January 2004 to December 2015, respectively) were used as the training sets to build the SARFIMA and SARIMA models, while the remaining periods served as the test sets to evaluate the forecasting accuracy of both models.

Research results

During the study period, a total of 23400874 HB cases and 3590867 HC cases were reported. In the 12-step-ahead HB, 90-step-ahead HB, 12-step-ahead HC, and 90-step-ahead HC forecasts, the best SARFIMA generated lower error rates compared with the best SARIMA. The predicted HB incidents totaled 9865400 [95% confidence interval (95%CI): 7508093-12222709] and HC totaled 1659485 (95%CI: 856681-2462290) during 2023-2030.

Research conclusions

The SARFIMA provides a more sophisticated and adaptable framework for capturing intricate patterns and interdependencies in monitoring HB and HC epidemics compared with the SARIMA. This ultimately leads to enhanced forecasting capabilities and a deeper comprehension of the underlying process.

Research perspectives

The integration of SARFIMA into public health decision-making for managing HB and HC epidemics can result in more informed and efficacious interventions.

ACKNOWLEDGEMENTS

We thank the Chinese CDC for sharing the HB and HC incidence series data in mainland China.

FOOTNOTES

Author contributions: Wang YB conceived, initiated, and performed this work; Qing SY, Ma C, Liang ZY, Bai YC, and Xu CJ collected, analyzed, and interpreted the data for this study; Wang YB, Qing SY, Liang ZY, Bai YC, Xu CJ, and Ma C edited and improved this original manuscript; all authors reviewed and approved the manuscript.

Supported by the Key Scientific Research Project of Universities in Henan Province, No. 21A330004; and Natural Science Foundation in Henan Province, No. 222300420265.

Institutional review board statement: This study was reviewed and approved by the institutional review board of Xinxiang Medical University (No: XYLL-2019072). All methods were carried out under relevant guidelines and regulations.

Informed consent statement: The need for informed consent was waived by the Ethics Committee of Xinxiang Medical University because the HB and HC cases were shared anonymously and we cannot access any identifying information of the patients (available from: <https://www.phsciencedata.cn/Share/>).

Conflict-of-interest statement: We have no financial relationships to disclose.

Data sharing statement: All the data supporting the findings of the work are contained within the study or technical appendix, statistical code, and dataset available from the corresponding author at 191035@xxmu.edu.cn.

Open-Access: This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <https://creativecommons.org/licenses/by-nc/4.0/>

Country/Territory of origin: China

ORCID number: Yong-Bin Wang 0000-0001-7854-7020; Si-Yu Qing 0009-0008-4916-5881; Zi-Yue Liang 0009-0009-7950-0545; Chang Ma 0009-0003-5940-5619; Yi-Chun Bai 0000-0003-0998-8441; Chun-Jie Xu 0009-0005-2824-1695.

S-Editor: Lin C

L-Editor: Wang TQ

P-Editor: Yu HG

REFERENCES

- 1 **World Health Organization.** Hepatitis. [cited 4 August 2023]. Available from: https://www.who.int/health-topics/hepatitis#tab=tab_1
- 2 **World Health Organization.** Global health sector strategy on viral hepatitis 2016-2021. Towards ending viral hepatitis. [cited 4 August 2023]. Available from: <https://www.who.int/publications/i/item/WHO-HIV-2016.06>
- 3 **GBD 2019 Hepatitis B Collaborators.** Global, regional, and national burden of hepatitis B, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Gastroenterol Hepatol* 2022; **7**: 796-829 [PMID: 35738290 DOI: 10.1016/S2468-1253(22)00124-8]
- 4 **Liu J, Liang W, Jing W, Liu M.** Countdown to 2030: eliminating hepatitis B disease, China. *Bull World Health Organ* 2019; **97**: 230-238 [PMID: 30992636 DOI: 10.2471/BLT.18.219469]
- 5 **Ceylan Z.** Estimation of COVID-19 prevalence in Italy, Spain, and France. *Sci Total Environ* 2020; **729**: 138817 [PMID: 32360907 DOI: 10.1016/j.scitotenv.2020.138817]
- 6 **Gan R, Chen N, Huang D.** Comparisons of forecasting for hepatitis in Guangxi Province, China by using three neural networks models. *PeerJ* 2016; **4**: e2684 [PMID: 27843718 DOI: 10.7717/peerj.2684]
- 7 **Wang YW, Shen ZZ, Jiang Y.** Comparison of ARIMA and GM(1,1) models for prediction of hepatitis B in China. *PLoS One* 2018; **13**: e0201987 [PMID: 30180159 DOI: 10.1371/journal.pone.0201987]
- 8 **Zheng Y, Zhang L, Zhu X, Guo G.** A comparative study of two methods to predict the incidence of hepatitis B in Guangxi, China. *PLoS One* 2020; **15**: e0234660 [PMID: 32579598 DOI: 10.1371/journal.pone.0234660]
- 9 **Qi C, Zhang D, Zhu Y, Liu L, Li C, Wang Z, Li X.** SARFIMA model prediction for infectious diseases: application to hemorrhagic fever with renal syndrome and comparing with SARIMA. *BMC Med Res Methodol* 2020; **20**: 243 [PMID: 32993517 DOI: 10.1186/s12874-020-01130-8]
- 10 **Tian CW, Wang H, Luo XM.** Time-series modelling and forecasting of hand, foot and mouth disease cases in China from 2008 to 2018. *Epidemiol Infect* 2019; **147**: e82 [PMID: 30868999 DOI: 10.1017/S095026881800362X]
- 11 **Siamba S, Otieno A, Koech J.** Application of ARIMA, and hybrid ARIMA Models in predicting and forecasting tuberculosis incidences among children in Homa Bay and Turkana Counties, Kenya. *PLOS Digit Health* 2023; **2**: e0000084 [PMID: 36812585 DOI: 10.1371/journal.pdig.0000084]
- 12 **Fallow, Sowell.** Modeling long-run behavior with the fractional ARIMA model. *J Monetary Eco* 1992; **29**: 277-302 [DOI: 10.1016/0304-3932(92)90016-U]
- 13 **Smith JP, Yadav S.** Forecasting costs incurred from unit differencing fractionally integrated processes. *Int J Forecasting* 1994; **10**: 507-514 [DOI: 10.1016/0169-2070(94)90019-1]
- 14 **Veenstra J.** Persistence and Anti-persistence: Theory and Software. PhD. Thesis, Western University. 2013
- 15 **Chang F, Huang H, Chan AHS, Shing Man S, Gong Y, Zhou H.** Capturing long-memory properties in road fatality rate series by an autoregressive fractionally integrated moving average model with generalized autoregressive conditional heteroscedasticity: A case study of Florida, the United States, 1975-2018. *J Safety Res* 2022; **81**: 216-224 [PMID: 35589293 DOI: 10.1016/j.jsr.2022.02.013]
- 16 **Chen S, Li J, Wang D, Fung H, Wong LY, Zhao L.** The hepatitis B epidemic in China should receive more attention. *Lancet* 2018; **391**: 1572 [PMID: 29695339 DOI: 10.1016/S0140-6736(18)30499-9]
- 17 **Bartholomew D, Box GEP, Jenkins GM, Ljung GM.** Time Series Analysis: Forecasting and Control. 5th ed. New Jersey: John Wiley and Sons, 2015: 14
- 18 **Hyndman RJ, Khandakar Y.** Automatic Time Series Forecasting: The forecast Package for R. *J Stat Softw* 2008; **27**: 1-22
- 19 **Dickey DP, Fuller WA.** Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* 1981; **49**: 1057-1071
- 20 **Alonso FJ, Pintado P, Del Castillo JM.** Filtering of kinematic signals using the Hodrick-Prescott filter. *J Appl Biomech* 2005; **21**: 271-285 [PMID: 16260847]
- 21 **Archibald BC, Koehler AB.** Normalization of seasonal factors in Winters' methods. *Int J Forecasting* 2003; **19**: 143-148 [DOI: 10.1016/S0169-2070(01)00117-0]
- 22 **Clegg LX, Hankey BF, Tiwari R, Feuer EJ, Edwards BK.** Estimating average annual per cent change in trend analysis. *Stat Med* 2009; **28**: 3670-3682 [PMID: 19856324 DOI: 10.1002/sim.3733]
- 23 **Belbute JM, Pereira AM.** Reference forecasts for CO(2) emissions from fossil-fuel combustion and cement production in Portugal. *Energy Policy* 2020; **144**: 111642 [PMID: 32565609 DOI: 10.1016/j.enpol.2020.111642]
- 24 **Boubaker H, Larbi OB.** Dynamic dependence and hedging strategies in BRICS stock markets with oil during crises. *Econ Anal Policy* 2022; **76**: 263-279 [DOI: 10.1016/j.eap.2022.08.011]
- 25 **Wang KW, Deng C, Li JP, Zhang YY, Li XY, Wu MC.** Hybrid methodology for tuberculosis incidence time-series forecasting based on ARIMA and a NAR neural network. *Epidemiol Infect* 2017; **145**: 1118-1129 [PMID: 28115032 DOI: 10.1017/S0950268816003216]
- 26 **Feroze N, Abbas K, Noor F, Ali A.** Analysis and forecasts for trends of COVID-19 in Pakistan using Bayesian models. *PeerJ* 2021; **9**: e11537 [PMID: 34277145 DOI: 10.7717/peerj.11537]
- 27 **Xiao Y, Li Y, Yu C, Bai Y, Wang L, Wang Y.** Estimating the Long-Term Epidemiological Trends and Seasonality of Hemorrhagic Fever with Renal Syndrome in China. *Infect Drug Resist* 2021; **14**: 3849-3862 [PMID: 34584428 DOI: 10.2147/IDR.S325787]
- 28 **Akhtar S, Rozi S.** An autoregressive integrated moving average model for short-term prediction of hepatitis C virus seropositivity among male volunteer blood donors in Karachi, Pakistan. *World J Gastroenterol* 2009; **15**: 1607-1612 [PMID: 19340903 DOI: 10.3748/wjg.15.1607]
- 29 **Wang Y, Wang M, Zhang G, Ou X, Ma H, You H, Jia J.** Control of Chronic Hepatitis B in China: Perspective of Diagnosis and Treatment. *China CDC Wkly* 2020; **2**: 596-600 [PMID: 34594716 DOI: 10.46234/ccdcw2020.159]

- 30 **Wang YS**, Wang SN, Pan JH, Wang WB. [Trend analysis and prediction of viral hepatitis incidence in China, 2009-2018]. *Zhonghua Liu Xing Bing Xue Za Zhi* 2020; **41**: 1460-1464 [PMID: 33076599 DOI: 10.3760/cma.j.cn112338-20191024-00761]
- 31 **Wang X**, Liu H, Qi J, Zeng F, Wang L, Yin P, Liu F, Li H, Liu Y, Liu J, Wei L, Liang X, Wang Y, Rao H, Zhou M. Trends of Mortality in End-Stage Liver Disease - China, 2008-2020. *China CDC Wkly* 2023; **5**: 657-663 [PMID: 37593126 DOI: 10.46234/ccdcw2023.128]
- 32 **Li J**, Ji XY, Geng J, Li N, Zhang GL, Zhao DY, Liu Y, Nie YG, Fan PY. Survey of prevalence of hepatitis C in people aged 1-69 years in Henan Province, 2020. *Zhonghua Liu Xing Bing Xue Za Zhi* 2023; **44**: 1114-1118 [PMID: 37482715 DOI: 10.3760/cma.j.cn112338-20220815-00711]
- 33 **Younossi ZM**, Tanaka A, Eguchi Y, Lim YS, Yu ML, Kawada N, Dan YY, Brooks-Rooney C, Negro F, Mondelli MU. The impact of hepatitis C virus outside the liver: Evidence from Asia. *Liver Int* 2017; **37**: 159-172 [PMID: 27748564 DOI: 10.1111/liv.13272]
- 34 **Mei X**, Lu H. Prevalence, diagnosis, and treatment of hepatitis C in Mainland China. *Glob Health Med* 2021; **3**: 270-275 [PMID: 34782868 DOI: 10.35772/ghm.2021.01080]
- 35 **Zhu Z**, Zhu X, Zhan Y, Gu L, Chen L, Li X. Development and comparison of predictive models for sexually transmitted diseases-AIDS, gonorrhea, and syphilis in China, 2011-2021. *Front Public Health* 2022; **10**: 966813 [PMID: 36091532 DOI: 10.3389/fpubh.2022.966813]
- 36 **Yu S**, Yu C, Li J, Liu S, Wang H, Deng M. Hepatitis B and hepatitis C prevalence among people living with HIV/AIDS in China: a systematic review and Meta-analysis. *Virology* 2020; **17**: 127 [PMID: 32831118 DOI: 10.1186/s12985-020-01404-z]
- 37 **Wang Y**, Jia J. Control of hepatitis B in China: prevention and treatment. *Expert Rev Anti Infect Ther* 2011; **9**: 21-25 [PMID: 21171874 DOI: 10.1586/eri.10.143]
- 38 **Luo T**, Lin Z, Wu Z, Cen P, Nong A, Huang R, Che J, Liang F, Yang Y, Liu J, Huang L, Cai J, Ou Y, Ye L, Bao L, Liang B, Liang H. Trends and associated factors of HIV, HCV and syphilis infection among different drug users in the China-Vietnam border area: an 11-year cross-sectional study (2010-2020). *BMC Infect Dis* 2023; **23**: 575 [PMID: 37667212 DOI: 10.1186/s12879-023-08239-3]



Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: <https://www.f6publishing.com/helpdesk>
<https://www.wjgnet.com>

