

# Validity of ICD-10-CM Codes Used to Identify Patients with Chronic Hepatitis B and C Virus Infection in Administrative Claims Data from the Taiwan National Health Insurance Outpatient Claims Dataset

This article was published in the following Dove Press journal:  
*Clinical Epidemiology*

Ming-Jen Sheu<sup>1,2</sup>  
Fu-Weng Liang<sup>3</sup>  
Sheng-Tun Li<sup>4</sup>  
Chung-Yi Li<sup>5-7</sup>  
Tsung-Hsueh Lu<sup>5</sup>

<sup>1</sup>Division of Gastroenterology and Hepatology, Chi Mei Medical Center, Tainan, Taiwan; <sup>2</sup>Department of Medicinal Chemistry, Chia Nan University of Pharmacy and Science, Tainan, Taiwan; <sup>3</sup>Department of Public Health, College of Health Science, Kaohsiung Medical University, Kaohsiung, Taiwan; <sup>4</sup>Department of Industrial and Information Management, College of Management, National Cheng Kung University, Tainan, Taiwan; <sup>5</sup>Department of Public Health, College of Medicine, National Cheng Kung University, Tainan, Taiwan; <sup>6</sup>Department of Healthcare Administration, College of Medical and Health Science, Asia University, Taichung, Taiwan; <sup>7</sup>Department of Public Health, College of Public Health, China Medical University, Taichung, Taiwan

**Purpose:** To validate the use of International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) codes to identify patients with chronic hepatitis B virus (HBV) and hepatitis C virus (HCV) infection in the Taiwan National Health Insurance (NHI) Outpatient Claims Dataset.

**Methods:** We conducted a retrospective study using results of HBV surface antigen (HBsAg), HBV e antigen (HBeAg), and anti-HCV antibody tests in the NHI Lab & Exam Dataset from January 1 to March 31, 2018, as the reference standard to confirm HBV and HCV infection cases. We calculated sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) to assess the performance of HBV infection-specific ICD-10-CM codes (B180, B181, and B191) and HCV infection-specific ICD-10-CM codes (B182 and B192) recorded in the NHI Outpatient Claims Dataset to identify patients with HBV or HCV infection.

**Results:** In total, 196,635 and 120,628 patients had analyzable results for HBsAg/HBeAg tests and anti-HCV tests, respectively. Moreover, 44,574 and 14,443 were confirmed to have HBV and HCV infection, respectively. The sensitivity, specificity, PPV, and NPV were, respectively, 46%, 83%, 45%, and 84% for HBV infection-specific ICD-10-CM codes and 47%, 99%, 81%, and 93% for HCV infection-specific ICD-10-CM codes. The sensitivity demonstrated great variation by region, clinical setting, and physician specialty.

**Conclusion:** The HBV and HCV infection-specific ICD-10-CM codes recorded by physicians in Taiwan NHI outpatient claims data in 2018 had moderate sensitivity and high specificity for both HBV and HCV infection. The PPV was high for HCV ICD-10-CM codes, yet moderate for HBV ICD-10-CM codes.

**Keywords:** chronic hepatitis B virus, chronic hepatitis C virus, international classification of diseases codes, validation, administrative claims data

Correspondence: Tsung-Hsueh Lu  
Department of Public Health, College of Medicine, National Cheng Kung University, No. 1, Dah Hsueh Road, East District, Tainan 701, Taiwan  
Tel +886 6 2353535 ext. 5567  
Fax +886 6 2359033  
Email robertlu@mail.ncku.edu.tw

## Introduction

Administrative claims data is a commonly used dataset for pharmacoepidemiological studies. Several scholars have used Taiwan National Health Insurance (NHI) claims data to examine the effects of drugs (e.g., antiviral, statin, and aspirin) among patients with chronic hepatitis B virus (HBV) or hepatitis C virus (HCV) infection.<sup>1-12</sup> As suggested by Chun et al, conducting a validation study to evaluate

the accuracy of coded algorithms to identify the health-related exposures, covariates, and outcomes is essential.<sup>13</sup> According to a systematic review of validation studies of claims data in the Asia-Pacific region, only 7 validation studies have been done in Taiwan, and all these studies focused on cardiovascular diseases.<sup>14</sup>

Regarding HBV and HCV, only 4 studies examined the validity of using International Classification of Diseases (ICD), Ninth Revision, Clinical Modification (ICD-9-CM) codes to identify patients with HBV or HCV infection.<sup>15–18</sup> The data used in 2 studies were confined to one medical center. These studies had small sample sizes ( $n = 331$  and  $200$ , respectively), and only the positive predictive value (PPV) was estimated to evaluate the validity of ICD-9-CM codes.<sup>15,16</sup> Two studies used data from 4 large health-care systems, had large sample sizes ( $n = 1,652,055$  and  $2,718,995$ , respectively), and calculated 4 performance indicators (i.e., sensitivity, specificity, PPV, and negative predictive value [NPV]) for HBV and HCV infection-related code algorithms.<sup>17,18</sup>

Chun et al further emphasized the need to investigate the transportability of coded algorithms to various populations in different health-care settings and especially when the coding system changed.<sup>13</sup> The ICD, Tenth Revision, Clinical Modification (ICD-10-CM) was introduced in the United States on October 1, 2015, and in Taiwan on January 1, 2016. Here, we validated the use of ICD-10-CM codes recorded in Taiwan NHI outpatient claims data to identify patients with HBV and HCV infections.

## Materials and Methods

### Study Setting

Taiwan has a population of 23 million. The single-payer NHI program, introduced in 1995, covers more than 99% of Taiwanese citizens. The NHI claims data (included outpatient and inpatient claims) are released to researchers as NHI Research Database (NHIRD). Since a change and update in 2016, NHIRD has been maintained and regulated by the Data Science Centre of the Ministry of Health and Welfare of Taiwan.<sup>19,20</sup>

The NHI began to reimburse antiviral therapy for people with HBV in 2003. As more effective antiviral drugs developed, the Taiwan Health Promotion Administration Ministry of Health and Welfare began offering free HBV and HCV screening tests for people aged 45 years or older. These tests have been performed since August 1, 2011, to identify patients with HBV or HCV infection so that they can receive antiviral

therapy to prevent the occurrence of liver cirrhosis and hepatocellular carcinoma.<sup>21</sup> Therefore, the number of people receiving HBV and HCV tests has increased considerably since then.

### Study Design and Data Source

We conducted a retrospective study using results in the NHI Lab & Exam Dataset as the reference standard and linked them to the NHI Outpatient Claims Dataset to assess the performance of ICD-10-CM codes in identifying patients with HBV and HCV infections. This study was approved by the Institute of Review Board of National Cheng Kung University Hospital with record number B-ER-107-394.

#### NHI Lab & Exam Dataset (Reference Standard)

To reduce the duplication of ordering the same test or examination for a given patient by different doctors in different clinical settings, the Taiwan NHI Administration created the MediCloud System in 2016.<sup>22</sup> A physician can query the MediCloud System to view the results of laboratory tests and reports of examinations (e.g., chest radiography, magnetic resonance imaging, and cardiac catheterization) performed in previous medical encounters to avoid duplicative testing and examination. The contracted clinics and hospitals have been required to submit the results of laboratory tests and reports of examinations to the NHI Lab & Exam Dataset since 2016.<sup>23</sup>

#### NHI Outpatient Claims Dataset (for ICD-10-CM Codes)

The NHI Outpatient Claims Dataset contains approximately 40 variables, including basic information of the patient (birthdate, sex, and patient identifier), date of visit, physician specialty, one primary ICD-10-CM diagnostic code, 4 secondary ICD-10-CM diagnostic codes, one primary procedure code, 2 secondary procedure codes, medication codes, test and examination codes, DRG code, and various reimbursement fees.<sup>24</sup> Here, we extracted information on patient age and sex, physician specialty, and ICD-10-CM codes for 5 diagnoses. In Taiwan, physicians record the ICD-10-CM codes of diagnoses and procedures in outpatient claims data.

### Study Population

We extracted results of HBV surface antigen (HBsAg; NHI order code 14032C), HBV e antigen (HBeAg; NHI order code 14035C), and anti-HCV antibody (NHI order code 14051C) tests submitted to the NHI Lab & Exam Dataset from January 1, 2018, to March 31, 2018, as the reference standard.

After we excluded missing or erroneous results (such as results of liver function tests rather than those of HBsAg or HBeAg tests), 267,397 HBsAg or HBeAg test results and 160,433 anti-HCV antibody results were analyzed (Figure 1).

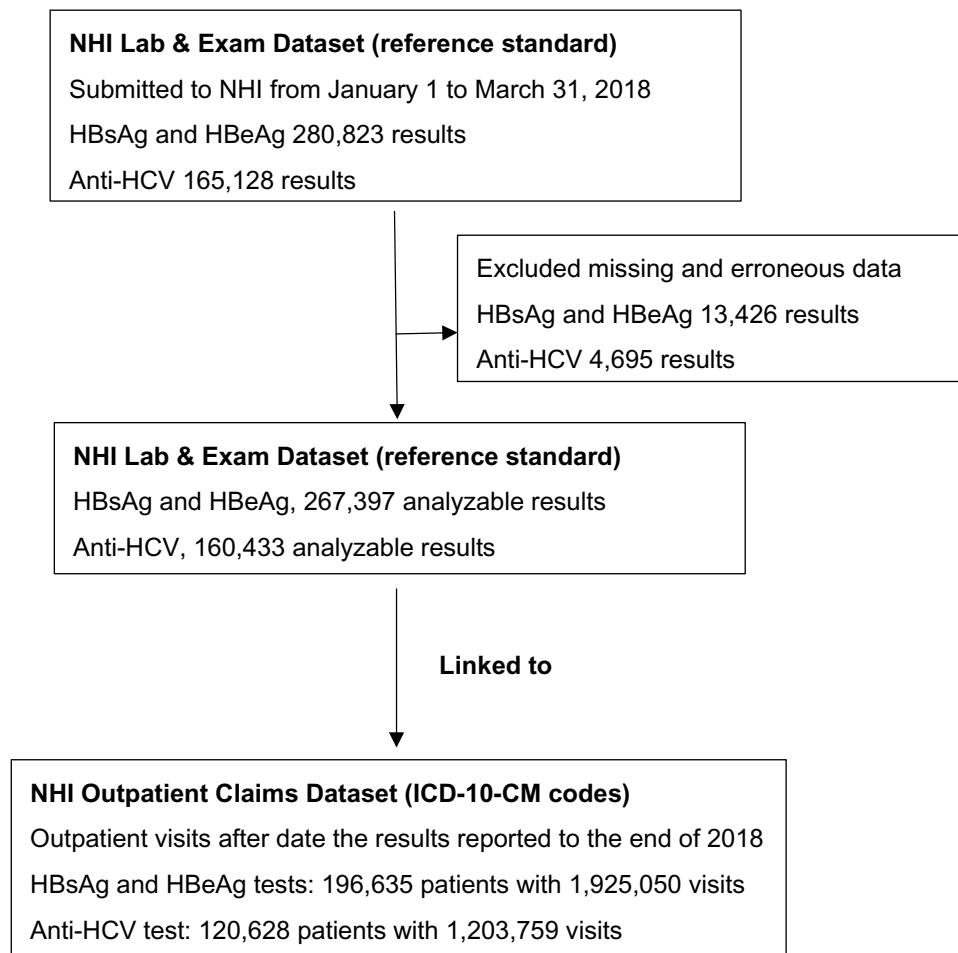
We then used unique personal identifiers to link these results to the NHI Outpatient Claims Dataset in 2018. We examined the ICD-10-CM codes recorded in all outpatient claims after the reporting date of test results to the end of 2018. Take two patients as example, if the result of an HBsAg test for patient A was reported on February 12, 2018, we examined the ICD-10-CM codes recorded in all outpatient claims from February 13 to December 31, 2018. In addition, if the result of an Anti-HCV test for patient B was reported on March 31, 2018, we examined the ICD-10-CM codes recorded in all outpatient claims visits from April 1 to December 31, 2018. That is to say that every patient had at least 9 months of observations. We assumed that the physicians would record HBV and HCV infection-specific ICD-10-CM codes after they noted that the HBV

or HCV infection-related test results were positive. We assumed that the physicians would record HBV and HCV infection-specific ICD-10-CM codes after they noted that the results of the HBV and HCV infection-related tests were positive.

The HBV infection-specific ICD-10-CM codes include B180 (chronic viral hepatitis B with delta agent), B181 (chronic viral hepatitis B without delta agent), and B191 (unspecified viral hepatitis B). The HCV infection-specific ICD-10-CM codes include B182 (chronic viral hepatitis C) and B192 (unspecified viral hepatitis C). We used only one coded algorithm in this study (i.e., once in outpatient claims data) due to its popular use in most studies employing NHI claims data for pharmacoepidemiological studies.<sup>1-12</sup>

## Statistical Analysis

To assess the performance of ICD-10-CM codes in identifying patients with HBV or HCV infection, the sensitivity, specificity, PPV, and NPV were calculated. In addition, we



**Figure 1** Flow of data linkage and exclusion.

computed the 95% confidence intervals (CI) for these 4 performance indicators.<sup>25</sup> The equation for 95% CI is

$$p \pm 1.96\sqrt{p(1-p)/n}$$

We also assessed the performance of ICD-10-CM codes by patient characteristics, regions, clinical setting, and physician specialty (gastroenterologists and hepatologists vs others).

## Results

### Study Population

For 267,397 HBV infection-related results, we linked to 1,925,050 outpatient visits for 196,635 patients. For 160,433 HCV infection-related results, we linked to 1,203,759 outpatient visits among 120,628 patients (Figure 1). The characteristics of patients who had analyzable results of HBsAg, HBeAg, and anti-HCV antibody tests and linked to NHI outpatient claims data, listed in Table 1. Compared with the general population in Taiwan, the patients who underwent testing for HBV and HCV infection were older.

### HBV Infection

Of 196,635 patients having HBsAg and HBeAg test results linked to NHI outpatient claims data, 44,574 (23%) had confirmed HBV, but only 20,621 (46%; i.e., sensitivity)

**Table 1** Characteristics of Subjects with Hepatitis B Virus Surface Antigen/Hepatitis B Virus e Antigen or Anti-Hepatitis C Virus Tests in Taiwan NHI Lab & Exam Dataset from January 1 to March 31, 2018, and in Taiwanese Population as a Whole

	HBsAg/HBeAg		Anti-HCV		Taiwan	
	No.	(%)	No.	(%)	No.	(%)
All patients	196,635	(100)	120,628	(100)	23,588,932	(100)
Sex						
Men	104,666	(53)	62,836	(52)	11,712,913	(50)
Women	91,968	(47)	57,791	(48)	11,876,019	(50)
Age						
≤44 years	50,994	(26)	30,274	(25)	13,097,880	(56)
45–64 years	95,386	(49)	54,440	(45)	7,057,535	(30)
≥65 years	48,755	(25)	35,914	(30)	3,433,517	(15)
Region						
Taipei	63,797	(32)	39,768	(33)	7,641,994	(32)
North	31,632	(16)	17,685	(15)	3,772,380	(16)
Central	31,678	(16)	18,784	(16)	4,578,749	(19)
South	31,904	(16)	20,646	(17)	3,345,543	(14)
Kauping	32,377	(16)	19,994	(17)	3,703,379	(16)
East	5,065	(3)	3,642	(3)	546,887	(2)

**Abbreviations:** HBsAg, hepatitis B virus surface antigen; HBeAg, hepatitis B virus e antigen; Anti-HCV, hepatitis C virus antibody.

patients were identified under the ICD-10-CM codes B180, B181, or B191. Of the 46,265 patients with HBV infection-specific ICD-10-CM codes, 20,621 were true positive and 25,644 were false positive, which resulted in a PPV of 45% (20,621 of 46,265; Table 2A). The specificity (i.e., ability of HBV infection-specific ICD-10-CM codes to identify those who did not have HBV infection correctly) was 83%.

Some of the performance indicators for HBV infection-specific ICD-10-CM codes showed great variations in some characteristics (Table 3). For instance, the sensitivity was the largest for the central region (59%), regional hospitals (54%), and gastroenterology specialty (66%), but lowest for the northern region (30%), clinics (17%), and specialties other than gastroenterology (21%). The PPV decreased with age, which was 51% for patients aged 44 years or younger and 40% for patients aged 65 years or older. No prominent differences by characteristics were noted for specificity and NPV.

### HCV Infection

Of 120,628 patients who had results of anti-HCV, 14,443 (12.0%) had confirmed HCV. The sensitivity (i.e., ability of HCV infection-specific ICD-10-CM codes B182 or B192 to correctly identify patients with HCV infection) was 47% (6,767 of 14,443). Of the 8,379 patients with HCV infection-specific ICD-10-CM codes, 6,767 were true positive and 1,612 were false positive, which resulted in a PPV of 81% (6,767 of 8,379; Table 2B). The specificity and NPV of HCV infection-specific ICD-10-CM codes were high: 99% (104,573 of 112,249) and 93% (104,573 of 112,249), respectively.

Similarly, we noted large variations in sensitivity across regions (54% in the southern region vs 32% in the eastern region), clinical settings (52% in regional hospitals vs 13% in clinics), and specialties (75% for gastroenterology vs 22% for other specialties). The specificity and NPV in different characteristics were similar and higher than 90% (Table 4). No age differences in PPV were found (79% for patients aged 44 years or younger and 82% for patients aged 65 years or older).

## Discussion

### Main Findings

The findings of this study indicate that both HBV and HCV infection-specific ICD-10-CM codes recorded by physicians in Taiwan NHI outpatient claims data have moderate sensitivity (46% and 47%, respectively). HCV

**Table 2** Performance of Using International Classification of Diseases, Tenth Revision, Clinical Modification, Codes in Identifying Patients with Chronic Hepatitis B (A) or C (B) Virus in Taiwan National Health Insurance Outpatient Claims Data, 2018

<b>A</b>			
ICD-10-CM Codes	Confirmed HBV Case	Not a HBV Case	Total
B180, B181, B191	20,621	25,644	46,265
No B180, B181, B191	23,953	126,417	150,370
Total	44,574	152,061	196,635
Sensitivity, 20,621/44,474 (46.3%) Specificity, 126,417/152,061 (83.1%) Positive predictive value, 20,621/46,265 (44.6%) Negative predictive value, 126,417/150,370 (84.1%)			
<b>B</b>			
ICD-10-CM Codes	Confirmed HCV Case	Not a HCV Case	Total
B182, B192	6,767	1,612	8,379
No B182, B192	7,676	104,573	112,249
Total	14,443	106,185	120,628
Sensitivity, 6,767/14,443 (46.9%) Specificity, 104,573/112,249 (98.5%) Positive predictive value, 6,767/8,379 (80.8%) Negative predictive value, 104,573/112,249 (93.2%)			

**Abbreviations:** HBV, chronic hepatitis B virus; HCV, chronic hepatitis C virus; ICD-10-CM, International Classification of Diseases Tenth Revision, Clinical Modification.

infection-specific ICD-10-CM codes had better performance than did HBV infection-specific ICD-10-CM codes in specificity (98% vs 83%), PPV (81% vs 45%), and NPV (93% vs 84%). The sensitivity varied greatly by region, clinical setting, and physician specialty. However, no prominent differences by characteristics were noted for specificity, PPV, and NPV. This valuable finding could help researchers design different code algorithms for different research questions on patients with chronic HBV or HCV.

## Comparison with Previous Studies

The PPV of HBV infection-specific ICD-10-CM codes in this study was 45%, similar to the PPV of ICD-9-CM codes in the Michael E. DeBakey Veterans Affairs Medical Center (MDVAMC) study (43%) and lower than that in the University of Pennsylvania Health System (UPHS) (81%) and the Centers for Disease Control and Prevention Chronic Hepatitis Cohort Study (CHeCS), including the data from four large health-care systems (61%).<sup>15–17</sup> The sensitivity and specificity for HBV

infection-specific codes were 46% and 83%, respectively, lower than in CHeCS (i.e., 84% and 99%, respectively).<sup>17</sup>

The PPV of HCV infection-specific ICD-10-CM codes was 81%, which was lower than that in the MDVAMC study (93%), UPHS (89%), and CHeCS (92%).<sup>15,16,18</sup> The sensitivity and specificity for HCV codes in this study were 47% and 99%, respectively—lower than those in CHeCS (i.e., 70% and 99%, respectively).<sup>18</sup>

The main reason for lower performance of ICD-10-CM codes in this study was the inclusion of reports from all contracted hospitals and clinics in Taiwan. The 2 studies with better performance in HBV and HCV infection identification were each confined to one medical center, with MDVAMC having more elderly patients with liver diseases and UPHS offering liver transplantation.<sup>15,16</sup> The findings of the present study also suggest better performance if the ICD-10-CM codes were recorded by gastroenterologists and hepatologists.

This population-based study results are relatively comparable to those of CHeCS, which included 2.7 million patients from 4 large health-care systems in different regions. A possible explanation of lower performance of HBV and HCV infection-specific codes in the present study are the quality of the reference standard used. We used only the results of laboratory tests submitted by contracted hospitals and clinics as a reference standard without reviewing the electronic medical record. However, the reference standard used in CHeCS included not only laboratory test results but also chart reviews.<sup>17,18</sup>

## Interpretation and Implications of Results

The moderate sensitivity of ICD-10-CM codes in identifying patients with HBV or HCV infection in this study may be explained by many HBV and HCV infection-related tests being performed as part of free adult preventive health checkups by hospitals where usual care is typically not provided. The reports of these health checkups were mailed to the patients themselves. The physician who provides usual care might not know the results of HBV and HCV-related tests if the patient does not bring the report to the consultation. The second possible explanation is the poor coding literacy for some physicians and not recording correct HBV and HCV infection-specific ICD-10-CM codes in the outpatient settings. Some physicians might have recorded ICD-10-CM code K73 “chronic hepatitis” instead.

To improve the use of ICD-10-CM codes for identifying patients with HBV or HCV infection, the NHI should



**Table 3** Performance of Using International Classification of Diseases, Tenth Revision, Clinical Modification, Codes (B180, B181, B191) to Identify Patients with Chronic Hepatitis B Virus in Taiwan National Health Insurance Outpatient Claims Data, by Characteristics, 2018

	Sen	(95% CI)	Spe	(95% CI)	PPV	(95% CI)	NPV	(95% CI)
All patients	46.3	(45.8–46.7)	83.1	(82.9–83.3)	44.6	(44.1–45.0)	84.1	(83.9–84.3)
Sex								
Men	49.9	(49.3–50.5)	79.5	(79.2–79.8)	43.9	(43.4–44.4)	83.1	(82.9–83.4)
Women	41.5	(40.8–42.2)	87.1	(86.9–87.3)	45.7	(45.0–46.5)	85.0	(84.8–85.3)
Age								
≤44 years	48.0	(47.1–48.9)	85.3	(85.0–85.7)	51.4	(50.5–52.3)	83.5	(83.2–83.9)
45–64 years	46.5	(45.9–47.1)	79.0	(78.7–79.3)	42.6	(42.0–43.2)	81.5	(81.2–81.8)
≥65 years	42.9	(41.8–44.0)	86.8	(86.5–87.2)	39.6	(38.6–40.6)	88.3	(88.0–88.6)
Region								
Taipei	47.5	(46.7–48.3)	83.2	(82.9–83.5)	46.0	(45.2–46.8)	84.0	(83.7–84.4)
North	29.6	(28.6–30.6)	90.9	(90.6–91.3)	54.2	(52.8–55.6)	78.1	(77.6–78.6)
Central	59.4	(58.2–60.5)	79.8	(79.3–80.3)	45.2	(44.2–46.2)	87.5	(87.1–87.9)
South	54.6	(53.3–56.0)	79.0	(78.5–79.5)	34.3	(33.3–35.3)	89.7	(89.3–90.0)
Kauping	46.8	(45.7–47.9)	83.2	(82.7–83.6)	46.8	(45.7–47.9)	83.2	(82.7–83.6)
East	32.0	(29.4–34.5)	85.1	(84.0–86.2)	42.2	(39.1–45.3)	78.6	(77.4–79.9)
Clinical setting								
Medical centers	46.8	(46.0–47.5)	82.1	(81.8–82.5)	51.1	(50.4–51.9)	79.4	(79.0–79.8)
Regional hospitals	53.8	(53.0–54.7)	79.8	(79.4–80.1)	39.1	(38.4–39.8)	87.8	(87.5–88.0)
District hospitals	47.6	(46.5–48.8)	85.0	(84.5–85.4)	44.7	(43.6–45.8)	86.4	(86.0–86.8)
Clinics	16.6	(15.4–17.7)	94.0	(93.7–94.4)	39.9	(37.6–42.2)	82.4	(81.9–82.9)
Physician specialty								
GI & Hepa	65.8	(65.2–66.4)	57.3	(56.9–57.8)	45.4	(44.9–45.9)	75.7	(75.2–76.1)
Others	21.1	(20.5–21.7)	94.5	(94.4–94.6)	41.5	(40.5–42.4)	86.6	(86.4–86.8)

**Abbreviations:** GI & Hepa, gastroenterology and hepatology; Sen, sensitivity; Spe, specificity; PPV, positive predictive value; NPV, negative predictive value.

encourage physicians to query the MediCloud System to confirm if the patient has HBV or HCV.<sup>22</sup> Furthermore, the NHI can provide feedback to or remind physicians if the recorded ICD-10-CM codes were not compatible (false-positive or false-negative) with the results of HBV or HCV laboratory tests.

The possible explanation for a better PPV associated with HCV-infection ICD-10-CM codes as compared to HBV-infection ICD-10-CM codes (81% vs 45%) was that the NHI began to reimburse HCV antiviral treatment since 2017 and the correct ICD-10-CM codes in outpatient claims data are one of the requirements for getting approval of antiviral treatment for patients with HCV.<sup>26,27</sup>

With regard to nearly half of patients with positive HBV-infection ICD-10-CM codes were false positive, the first possible explanation was that many physicians assigned the HBV-infection related codes to avoid the denial of reimbursement. The physicians did not check the HBV-related test results, and still copied-and-pasted the same diagnosis even after the test results were available. The second possible

explanation was that physicians might have other evidence of HBV-infection other than the test results performed during the first three months of 2018 such as results from private paid tests. The third possible explanation was that we used a less strict algorithm, i.e., one-time HBV-infection ICD-10-CM code at outpatient claims for case identification, which was subject to increase the false positive rate. A lower false positive rate can be expected if we used a stricter algorithm (e.g., considering at least three or more outpatient visits with HBV-infection ICD-10-CM code as positive diagnosis).

Further studies are needed to clarify the specific reasons that account for the low PPV for using HBV-infection ICD-10-CM codes to identify cases. Scholars in future studies could examine the outpatient diagnosis codes and other related evidences before 2018 for patients been judged as false positive HBV-infection ICD-10-CM codes in this study. Hospital-based studies are also needed to obtain laboratory tests results and more detail medical record for longer years to establish more robust reference standard to assess the performance of outpatient HBV-related ICD-10-CM codes.

**Table 4** Performance of Using International Classification of Diseases, Tenth Revision, Clinical Modification, Codes (B182, B192) to Identify Patients with Chronic Hepatitis C Virus in Taiwan National Health Insurance Outpatient Claims Data, by Characteristics, 2018

	Sen	(95% CI)	Spe	(95% CI)	PPV	(95% CI)	NPV	(95% CI)
All patients	46.9	(46.0–47.7)	98.5	(98.4–98.6)	80.8	(79.9–81.6)	93.2	(93.0–93.3)
Sex								
Men	47.8	(46.6–48.9)	98.6	(98.5–98.7)	80.9	(79.7–82.0)	93.6	(93.4–93.8)
Women	46.0	(44.8–47.1)	98.4	(98.3–98.5)	80.7	(79.5–81.9)	92.6	(92.4–92.9)
Age								
≤44 years	47.2	(44.8–49.6)	99.2	(99.1–99.3)	78.6	(76.0–81.1)	97.0	(96.8–97.2)
45–64 years	47.9	(46.7–49.1)	98.4	(98.3–98.5)	80.2	(78.9–81.4)	93.3	(93.1–93.5)
≥65 years	45.7	(44.4–46.9)	97.9	(97.7–98.0)	82.0	(80.7–83.3)	89.5	(89.2–89.8)
Region								
Taipei	48.4	(46.7–50.1)	98.4	(98.3–98.6)	73.9	(72.0–75.7)	95.5	(95.3–95.7)
North	36.1	(34.0–38.3)	99.3	(99.2–99.4)	86.2	(83.8–88.5)	92.8	(92.4–93.1)
Central	46.8	(44.5–49.1)	98.8	(98.7–99.0)	80.7	(78.3–83.1)	94.6	(94.3–94.9)
South	54.3	(52.9–55.7)	98.6	(98.4–98.8)	92.1	(91.1–93.1)	87.8	(87.3–88.3)
Kauping	40.3	(38.3–42.3)	97.6	(97.4–97.8)	68.9	(66.5–71.4)	92.6	(92.2–92.9)
East	32.4	(27.4–37.4)	97.3	(96.7–97.8)	54.8	(47.9–61.7)	93.4	(92.6–94.2)
Clinical setting								
Medical centers	48.2	(46.6–49.7)	98.8	(98.7–99.0)	83.2	(81.7–84.7)	94.2	(93.9–94.4)
Regional hospitals	52.0	(50.8–53.2)	97.8	(97.6–97.9)	79.9	(78.7–81.1)	92.3	(92.0–92.6)
District hospitals	43.3	(41.5–45.1)	98.9	(98.7–99.0)	85.4	(83.6–87.3)	92.1	(91.7–92.5)
Clinics	13.2	(11.0–15.5)	99.0	(98.8–99.2)	46.0	(39.8–52.2)	94.7	(94.3–95.1)
Physician specialty								
GI & Hepa	75.2	(74.2–76.2)	94.1	(93.8–94.4)	81.7	(80.8–82.7)	91.6	(91.2–91.9)
Others	21.9	(20.9–22.8)	99.5	(99.4–99.5)	78.0	(76.2–79.7)	93.5	(93.3–93.7)

**Abbreviations:** GI & Hepa, gastroenterology and hepatology; Sen, sensitivity; Spe, specificity; PPV, positive predictive value; NPV, negative predictive value.

## Strengths and Limitations

This is the first validation study in the literature on HBV and HCV infection ICD-10-CM codes in administrative data at the nationwide level. Nevertheless, it has several limitations. First, we used only HBV and HCV test results in the first 3 months of 2018 and assessed the quality of coding after the results reported in 2018, the findings might be different from the quality of coding before 2018. Second, it is possible, that a patient received more accurate serological tests after March, 2018, and had a positive test result different from that done during the first 3 months. As such, the physicians might later assign different (more accurate) ICD-10-CM codes in this patient's outpatient claims. This patient was therefore mistakenly judged as “false positive”. However, the chance of such scenario is small because physicians are not likely to repeatedly order the serologic test within such a short period of time. Third, we used only HBV and HCV-related laboratory tests as a reference standard without reviewing other information recorded in the electronic medical record. Fourth, the

quality of laboratory tests in different hospitals and clinics might have differed. Finally, we examined only one coded algorithm (i.e., at least one outpatient claims with HBV and HCV infection-specific ICD-10-CM codes). Additional studies comparing the performance of different-coded algorithms, such as at least two ICD-10-CM codes and at least two ICD-10-CM codes separated by 6 months, are needed.

## Conclusion

According to the findings of this study, the HBV and HCV infection-specific ICD-10-CM codes recorded by physicians in Taiwan NHI outpatient claims data in 2018 had moderate sensitivity and high specificity. With regard to PPV, it is high for HCV ICD-10-CM codes, yet moderate for HBV ICD-10-CM codes. Further validation studies using longer period and more information for reference standard and strict algorithms are needed to provide more practical information for users to better identify patients with HBV or HCV infection using outpatient diagnostic codes.

## Disclosure

The authors report no conflicts of interest in this work.

## References

1. Wu CY, Chen YJ, Ho HJ, et al. Association between nucleoside analogues and risk of hepatitis B virus-related hepatocellular carcinoma recurrence following liver resection. *JAMA*. 2012;308(18):1906–1914. doi:10.1001/2012.jama.11975
2. Tsan YT, Lee CH, Wang JD, Chen PC. Statins and the risk of hepatocellular carcinoma in patients with hepatitis B virus infection. *J Clin Oncol*. 2012;30(6):623–630. doi:10.1200/JCO.2011.36.0917
3. Tsan YT, Lee CH, Ho WC, et al. Statins and the risk of hepatocellular carcinoma in patients with hepatitis C virus infection. *J Clin Oncol*. 2013;31(12):1514–1521. doi:10.1200/JCO.2012.44.6831
4. Hsu YC, Ho HJ, Wu MS, Lin JT, Wu CY. Postoperative peg-interferon plus ribavirin is associated with reduced recurrence of hepatitis C virus-related hepatocellular carcinoma. *Hepatology*. 2013;58(1):150–157. doi:10.1002/hep.26300
5. Hsu YC, Lin JT, Ho HJ, et al. Antiviral treatment for hepatitis C virus infection is associated with improved renal and cardiovascular outcomes in diabetic patients. *Hepatology*. 2014;59(4):1293–1302. doi:10.1002/hep.26892
6. Chiang CJ, Yang YW, Chen JD, et al. Significant reduction in end-stage liver diseases burden through the national viral hepatitis therapy program in Taiwan. *Hepatology*. 2015;61(4):1154–1162. doi:10.1002/hep.27630
7. Hsu YC, Ho HJ, Huang YT, et al. Association between antiviral treatment and extrahepatic outcomes in patients with hepatitis C virus infection. *Gut*. 2015;64(3):495–503. doi:10.1136/gutjnl-2014-308163
8. Yang YH, Chen WC, Tsan YT, et al. Statin use and the risk of cirrhosis development in patients with hepatitis C virus infection. *J Hepatol*. 2015;63(5):1111–1117. doi:10.1016/j.jhep.2015.07.006
9. Huang YW, Lee CL, Yang SS, et al. Statins reduce the risk of cirrhosis and its decompensation in chronic hepatitis B patients: a nationwide cohort study. *Am J Gastroenterol*. 2016;111(7):976–985. doi:10.1038/ajg.2016.179
10. Lee TY, Lin JT, Zeng YS, et al. Association between nucleos(t)ide analog and tumor recurrence in hepatitis B virus-related hepatocellular carcinoma after radiofrequency ablation. *Hepatology*. 2016;63(5):1517–1527. doi:10.1002/hep.28266
11. Chang FM, Wang YP, Lang HC, et al. Statins decrease the risk of decompensation in hepatitis B virus- and hepatitis C virus-related cirrhosis: a population-based study. *Hepatology*. 2017;66(3):896–907. doi:10.1002/hep.29172
12. Lee TY, Hsu YC, Tseng HC, et al. Association of daily aspirin therapy with risk of hepatocellular carcinoma in patients with chronic hepatitis B. *JAMA Intern Med*. 2019;179(5):633–640. doi:10.1001/jamainternmed.2018.8342
13. Chun DS, Lund JL, Stürmer T. Pharmacoepidemiology and drug safety's special issue on validation studies. *Pharmacoepidemiol Drug Saf*. 2019;28(2):123–125. doi:10.1002/pds.4694
14. Koram N, Delgado M, Stark JH, Setoguchi S, de Luise C. Validation studies of claims data in the Asia-Pacific region: a comprehensive review. *Pharmacoepidemiol Drug Saf*. 2019;28(2):156–170. doi:10.1002/pds.4616
15. Kramer JR, Davila JA, Miller ED, et al. The validity of viral hepatitis and chronic liver disease diagnoses in Veterans Affairs administrative databases. *Aliment Pharmacol Ther*. 2008;27(3):274–282. doi:10.1111/j.1365-2036.2007.03572.x
16. Niu B, Forde KA, Goldberg DS. Coding algorithms for identifying patients with cirrhosis and hepatitis B or C virus using administrative data. *Pharmacoepidemiol Drug Saf*. 2015;24(1):107–111. doi:10.1002/pds.v24.1
17. Mahajan R, Moorman AC, Liu SJ, et al. Use of the international classification of diseases, 9th revision, coding in identifying chronic hepatitis B virus infection in health system data: implications for national surveillance. *J Am Med Inform Assoc*. 2013;20(3):441–445. doi:10.1136/amiajnl-2012-001558
18. Abara WE, Moorman AC, Zhong Y, et al. The predictive value of international classification of disease codes for chronic hepatitis C virus infection surveillance: the utility and limitations of electronic health records. *Popul Health Manag*. 2018;21(2):110–115. doi:10.1089/pop.2017.0004
19. Hsing AW, Ioannidis JPA. Nationwide population science: lessons from the Taiwan national health insurance research database. *JAMA Intern Med*. 2015;175(9):1527–1529. doi:10.1001/jamainternmed.2015.3540
20. Lin LY, Warren-Gash C, Smeeth L, Chen PC. Data resource profile: the National Health Insurance Research Database (NHIRD). *Epidemiol Health*. 2018;40:e2018062. doi:10.4178/epih.e2018062
21. Taiwan Health Promotion Administration. Criteria for HBV and HCV tests in adult preventive health checkup. Available from <https://www.hpa.gov.tw/Pages/Detail.aspx?nodeid=814&pid=4496>. Accessed Aug 30, 2019.
22. Taiwan National Health Insurance Administration. Introduction of National Health Insurance MediCloud System. Available from [https://www.nhi.gov.tw/Content\\_List.aspx?n=B5612D76EB95D83D&topn=CA428784F9ED78C9](https://www.nhi.gov.tw/Content_List.aspx?n=B5612D76EB95D83D&topn=CA428784F9ED78C9). Accessed Aug 30, 2019.
23. Taiwan National Health Insurance Administration. Guidelines for submission of results of laboratory tests and reports of examinations. Available from [https://www.nhi.gov.tw/Content\\_List.aspx?n=264416706E2EF4DA&topn=D39E2B72B0BDF15](https://www.nhi.gov.tw/Content_List.aspx?n=264416706E2EF4DA&topn=D39E2B72B0BDF15). Accessed Aug 30, 2019.
24. Taiwan National Health Insurance Administration. Claims format for contracted clinics and hospitals. Available from [https://www.nhi.gov.tw/Content\\_List.aspx?n=AAA5E2B2776DD4CD&topn=D39E2B72B0BDF15](https://www.nhi.gov.tw/Content_List.aspx?n=AAA5E2B2776DD4CD&topn=D39E2B72B0BDF15). Accessed Aug 30, 2019.
25. Altman D, Machin D, Bryant T, Gardner M. *Statistics with Confidence: Confidence Intervals and Statistical Guidelines*. 2nd ed. London: BMJ Books; 2000.
26. Chen DS. Taiwan commits to eliminating hepatitis C in 2025. *Lancet Inf Dis*. 2019;19(5):466–467. doi:10.1016/S1473-3099(19)30170-7
27. Liu CH, Yu ML, Peng CY, et al. Real-world anti-viral treatment decisions among chronic hepatitis C patients in Taiwan: the INITIATE study. *J Formos Med Assoc*. 2019;118(6):1014–1023. doi:10.1016/j.jfma.2018.10.020

### Clinical Epidemiology

### Publish your work in this journal

Clinical Epidemiology is an international, peer-reviewed, open access, online journal focusing on disease and drug epidemiology, identification of risk factors and screening procedures to develop optimal preventative initiatives and programs. Specific topics include: diagnosis, prognosis, treatment, screening, prevention, risk factor modification,

Submit your manuscript here: <https://www.dovepress.com/clinical-epidemiology-journal>

systematic reviews, risk & safety of medical interventions, epidemiology & biostatistical methods, and evaluation of guidelines, translational medicine, health policies & economic evaluations. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use.

Dovepress