Journal of Korean Medical Science
Vol 5, No. 4, 179-183 December 1990

# Incidence Rate of Hepatitis B: Results of a Record Linkage Study Among Healthy Korean Males+

Keun-Young Yoo, M.D., Byung-Joo Park, M.D., Hyung-Sik Ahn, M.D., Dae-Hee Kang, M.D., Yoon-Ok Ahn, M.D., Hyo-Suk Lee, M.D.* and Chung-Yong Kim, M.D.*

*Department of Preventive Medicine and Department of Internal Medicine*, Seoul National University College of Medicine, Seoul, Korea*

*A stable, liver disease-free cohort of 369,725 was reconstructed as of 1984 for the study of hepatitis B virus seromarkers and liver diseases. The cohort consisted of male beneficiaries of the Korea Medical Insurance Corporation (KMIC) over 30 years of age and living nationwide. Subjects who were both negatives for HBsAg and anti-HBs (N=274,037) were selected for incidence of hepatitis B. Data on test results of HBsAg and anti-HBs in 1984 and on hepatitis B occurrence during 1985-1986 were collected from the files of the KMIC. Linkage was done between these two data sets to measure the incidence rate through a longitudinal observation of the male population. Correction against misclassification error and duplicate claims was done by a sample survey and verification procedures. The incidence rate of hepatitis B was 17.13 per 100,000 person-years for acute viral hepatitis B and 15.74 for chronic hepatitis B, respectively.*

*An increasing age-dependent pattern for acute hepatitis B was not so prominent in this population. However, the incidence rate of chronic hepatitis B steadily increased with age. The relative risk, estimated by a log-linear model for rate and constant hazard, was significantly higher in the over-60 age group than in the others. The incidence rate in the lower socioeconomic class was higher than in the others, although statistically not significant.*

Key Words: *Hepatitis B, Acute hepatitis B, Chronic hepatitis B, Incidence, Record linkage study*

## INTRODUCTION

Korea is known to be one of the most common endemic areas of the world for hepatitis B virus (HBV) infection. Chronic liver diseases such as cirrhosis and hepatocellular carcinoma are one of major health problems in this country. A recent report revealed that almost 85% of the male and over 60% of the female adult population had been infectecd with HBV when the infection was defined as positive for at least one seromarker among HBsAg, anti-HBs, and anti-HBc by radioimmunoassay (RIA) method (Yoo et al., 1988). In addition to that, it has been suggested that such a high level of infection of HBV might be largely due to the nonparenteral transmission of the virus, as well as to frequent infection from the relatively large pool of chronic carriers in the population; the point prevalence of HBsAg positivity was estimated as 11.7% in males and 9.5% in females (Ahn and Yoo, 1983; Ahn et al., 1987; Yoo et al., 1988).

Recently, academic and administrative interest has centered on the incidence of hepatitis B, instead of viral infection. The purpose of this study was to measure the incidence rate of hepatitis B through a longitudinal observation in a liver disease-free male cohort in Korea, as well as to observe the incidence rate by sociodemographic characteristics.

Address correspondence to: *Keun-Yong Yoo, M.D., Department of Preventive Medicine, Seoul National University College of Medicine, 28 Yongon-dong, Chongno-gu, Seoul 110-460, Korea (Tel: 02-760-3303)*

*+This work was supported by the Indirect Research Fund of Korea Research Foundation (1987).*

## MATERIALS AND METHODS

The study population was drawn from the beneficiaries of the Korea Medical Insurance Corporation (KMIC), consisting of governmental employees, private schoolteachers and staffs, and their pensioners from all the provinces. The eligible population for the cohort of male over 30 years of age was 516,668 persons in 1984. Excluded were physically unhealthy persons who were in need of further medical attention by physicians, those with abnormal liver function, those who had vaccination against HBV prior to recruitment, those without records on HBV markers, and those who migrated into or out of the study population during the observation period of 1984-1986. A stable and liver disease-free study population was reconstructed as of 1984. The cohort for the study on HBV seromarkers and liver diseases numbered 369,725.

Data on hepatitis B occurrence were collected from files stored in the main computer of the KMIC. When an insurance claim for medical service was reimbursed, a computerized record was entered into an automated "Treatment File" which contained information on the identification number and name of the patient, a three-digit code for the International Classification of Diseases (ICD) of the two presumptive diagnoses, the date of admission, and the names of hospitals and clinics, etc. Among the records submitted to the KMIC during January 1, 1985 to December 31, 1986 for any medical reason, 2,657,719 records were collected from the "Treatment File."

Linking was done between the "Treatment File" and the "Physical Examination File" of 1984; the later contained records of individual's HBV seromarkers by reversed passive hemagglutination (RPHA) for HBsAg and by passive hemagglutination (PHA) for anti-HBs. In the data linkage, an identification number was used. From the linked files, subjects who were both negatives for HBsAg and anti-HBs (n=274,037) were selected for incidence of hepatitis B. Admitted cases with a diagnosis of ICD-070 (acute hepatitis B), ICD-571 (chronic hepatitis B including chronic active hepatitis, chronic persistent hepatitis and liver cirrhosis) and ICD-155 (liver cancer) were measured and regarded as presumptive hepatitis cases.

In order to sort out the true hepatitis B from the presumptive cases and to unify the diagnostic criteria of AVH and CVH, a sample survey was carried out from June to July, 1988. Eight hospitals located in Seoul were chosen for the sample survey. Seventeen senior students of Seoul National University College of Medicine visited every hospital in which the presumptive cases had once been admitted. They abstracted 572 medical records using a standard dictation sheet. All medical records abstracted were reviewed by a specialist of hepatology. Acute and chronic hepatitis were confirmed by the criteria of most probable diagnosis.

Viral hepatitis was defined by several criteria: HBV seromarkers done by any kind of detection method, laboratory findings, including liver function test, and clinical findings such as hepatomegaly and jaundice. Differentiation between acute and chronic hepatitis was determined on the basis of the history of the present illness, regardless of the IgM-or IgG-component of antibody to core antigen.

The incidence rate was measured from the estimated number of acute hepatitis B or chronic hepatitis B/100,000 person-years. The actual number reported to the KMIC was corrected by various coefficients against misclassification errors between the true acute hepatitis B (or chronic hepatitis B) and ICD-070 ($\alpha$), ICD-571 ($\beta$) and ICD-155 ($\gamma$). Correction against duplicate claims ($E_1$, $E_2$, $E_3$) was also done for a valid estimation of incidence. Correction coefficients were obtained from the sample survey described above. The values and calculation procedures for correction were as follows:

Incidence rate for acute (or chronic) hepatitis B:

$$\frac{[C_{070} \times \alpha \times E_1] + [C_{571} \times \beta \times E_2] + [C_{155} \times \gamma \times E_3]}{Ni \times k} \times 100{,}000$$

- $N_i$ : number of cohort observed
- k : No. of years observed (=2)
- $C_{070}$ : number of claims with a presumptive diagnosis of ICD-070
- $C_{571}$ : number of claims with a presumptive diagnosis of ICD-571
- $C_{155}$ : number of claims with a presumptive diagnosis of ICD-155
- $\alpha$ : proportion of most probable cases among $C_{070}$ (0.348 for acute hepatitis B; 0.261 for chronic hepatitis B)
- $\beta$ : proportion of most probable cases among $C_{571}$ (0.018 for acute hepatitis B; 0.256 for chronic hepatitis B)
- $\gamma$ : proportion of most probable cases among $C_{155}$ (0.000 for acute hepatitis B; 0.011 for chronic hepatitis B)
  hepatitis B)
- $E_1$ : ratio of the number of admissions to the num-

ber of the claims with $C_{070}$ (=1.000)

$E_2$ : ratio of the number of admissions to the number of the claims with $C_{571}$ (=0.813)

$E_3$ : ratio of the number of admissions to the number of the claims with $C_{155}$ (=0.659)

Given the incidence rate of the lowest stratum of each sociodemographic variable as a reference level, the relative risk in each stratum was measured from regression coefficient estimated by maximum likelihood method in log-linear model, which was developed for analysis of rate and constant hazard in a followup study (Holford, 1980). The Wald statistic $Z=[\beta/S.E.]^2$ was used to test the statistical significance between the relative risk of each stratum. A likelihood ratio test using log likelihood statistics $\Delta G^2$ was performed to test the global null hypothesis of $H_0$: $\beta_1=\beta_2=\ldots=\beta_i=0$, and to test the linear trend of the variable (Breslow and Day, 1980; Holford, 1984). The SAS and the GLIM were computer systems used for this statistical analysis (SAS Institute, 1987; The GLIM Working Party, 1987). Socioeconomic status was classified into three levels by standard monthly wages.

## RESULTS

The incidence rate of hepatitis B among healthy Korean males was 17.13 per 100,000 person-years for acute hepatitis B and 15.74 for chronic hepatitis B, respectively (Table 1). An age-dependent increasing pattern for acute hepatitis B fluctuated and was statistically not significant in this adult population ($\Delta G^2=0.15$, df=3, ns). The incidence rate of chronic hepatitis B steadily increased with age, although there was no significant linear trend ($X^2$ for linear trend=2.35, df=1, ns) and no significant difference among age strata ($\Delta G^2=2.55$, df=3, ns). However, the relative risk of the age group over 60 was estimated as 1.83 by log-linear model, which was significantly higher than the reference level ($W^2=2.13$, $p<0.05$).

The incidence rate of acute hepatitis B seemed to be inversely associated with socioeconomic class (Table 2). The incidence rate in the lower socioeconomic class was higher than the others, which was statistically not significant, however ($\Delta G^2=0.16$, df=2, ns). The incidence rate of chronic hepatitis B was also not affected by socioeconomic status ($\Delta^2=0.23$, df=2, ns).

**Table 1.** Age-specific incidence rate of hepatitis B among HBsAg-antiHBs negatives[1] from stable male cohorts of a Korea Medical Insurance Corporation reconstructed as of 1984 in Korea, 1985-1986.

| Age Group | No. of Cohorts[2] | No. of Cases[3] | I.R.[4] | I.R.[5] | W[2 6] |
|---|---|---|---|---|---|
| Acute Hepatitis B | | | | | |
| 30-39 | 60,359 | 20.5 | 17.01 | 1.00 | - |
| 40-49 | 109,435 | 36.4 | 16.64 | 0.98 | 0.01 |
| 50-59 | 84,741 | 29.4 | 17.32 | 1.02 | 0.00 |
| 60- | 19,502 | 7.6 | 19.38 | 1.14 | 0.10 |
| Total | 274,037 | 93.97 | 17.13 | $\Delta G^{2\,7}=0.15(3)$ | |
| Chronic Hepatitis B | | | | | |
| 30-39 | 60,359 | 15.8 | 13.08 | 1.00 | - |
| 40-49 | 109,435 | 31.6 | 14.43 | 1.10 | 0.10 |
| 560-59 | 84,741 | 29.6 | 17.44 | 1.33 | 0.86 |
| 60- | 19,502 | 9.3 | 23.89 | 1.83 | 2.13 |
| Total | 274,037 | 86.3 | 15.74 | $\Delta G^{2\,7}=2.55(3)$ | |

1. Tested by RPHA for HBsAg, and by PHA for anti-HBs at the time of recruitment (1984)
2. Number of cohorts at the time of recruitment
3. Estimated number of clinical cases during 1986-1987 from the calculation by correction coefficients for misclassification of ICD codes and for duplicate claims
4. Incidence rate per 100,000 person-years
5. Relative risk calculated from regression coefficient ß estimated by maximum likelihood method in log linear model for rate and constant hazard
6. Wald statistic: $Z=[\beta/S.E.]^2$ for $H_0$: $\beta\,(SES)_i=0$
7. Log likelihood statistics: $\Delta G^2$ for $H_0$: $\beta_1=\beta_2=\ldots=\beta_i=0$

## DISCUSSION

Korea, like many other countries in the Asia-Pacific region, has an important health problem with hepatitis, not only because of its high prevalence, but also because of its poor prognosis related to cirrhsis or hepatocellular carcinoma (Blumberg et al., 1975; Szmuness, 1978; Beasley et al., 1981). Most endeavors against this hepatic problem have concentrated on vaccination. A hepatitis B vaccine was developed in Korea (Kim, 1979), and since 1983 the government adopted a vaccination program against HBV. The results ot this study are sure to be of value not only to understand the natural history of HBV, but also to establish a national health plan against liver diseases.

Every person insured by KMIC is eligible to be covered for 180 hospitaldays annually according to the insurance policy. Under such circumstances, several inherent problems, such as duplicate counting, diag-

**Table 2.** Incidence rate of hepatitis B by socioeconomic status among HBsAg-antiHBs negatives[1] from stable male cohorts of the Korea Medical Insurance Corporation reconstructed as of 1984 in Korean, 1985-1986

| S.E.S.[2] | No. of Cohorts[3] | No. of Cases[4] | I.R.[5] | R.R.[6] | W[2 7] |
|---|---|---|---|---|---|
| Acute Hepatitis B | | | | | |
| Lower | 69,201 | 24.73 | 17.87 | 1.00 | |
| Middle | 145,843 | 50.40 | 17.28 | 0.97 | 0.02 |
| Upper | 58,993 | 18.74 | 15.88 | 0.89 | 0.15 |
| Total | 274,037 | 93.87 | 17.13 | $\Delta G^{2\,8}=0.16(2)$ | |
| Chronic Hepatitis B | | | | | |
| Lower | 69,201 | 22.79 | 16.47 | 1.00 | |
| Middle | 145,843 | 43.58 | 14.94 | 0.91 | 0.14 |
| Upper | 58,993 | 16.82 | 14.26 | 0.87 | 0.20 |
| Total | 274,037 | 86.25 | 15.74 | $\Delta G^{2\,8}=0.23(2)$ | |

1. Tested by RPHA for HBsAg, and by PHA for anti-HBs at the time of recruitment (1984)
2. Socioeconomic status classified by standard monthly wages
3. Number of cohorts at the time of recruitment
4. Estimated number of clinical cases during 1986-1987 from the calculation by correction coefficients for misclassification of ICD codes and for duplicate claims
5. Incidence rate per 100,000 person-years
6. Relative risk calculated from regression coefficient ß estimated by maximum likelihood method in log-linear model for rate and constant hazard $=0$
7. Wald statistic: $t=[ß/S.E.]^2$ for $H_0$: ß (SES) $=0$
8. Log likelihood statistics: $\Delta G^2$ for $H_0$: $ß_1=ß_2=\ldots=ß_i$

nostic inaccuracy, and misclassification errors in reporting, are unavoidable, which might adversely affect research purposes. In order to get a valid estimate of the health index (especially morbidity) in utilizing such insurance data, a verification procedure should be applied. Though correction coefficients from the sample survey used in this study may not exactly reflect the parameters in the population, such a procedure is the best way to come close to the true value.

The RPHA-PHA method is known to be less valid than the RIA method, especially in terms of specificity (Kim et al., 1984; Park, 1987). This is another source of invalidity in estimating true incidence in this population. In order to get a valid estimate of the incidence rate, one may consider application of predictability of RPHA-PHA for HBV infection. Theoretically, one must apply a predictive value drawn from the very same population. In order to get a valid estimate of the incidence rate, one may consider application of predictability of RPHA-PHA for HBV infection. Theoretically, one must apply a predictive value drawn from the very same population. The positive and negative predictive value of RPHA-PHA for the prediction of true positive by radioimmunoassay (RIA: AUSRIA-AUSAB -CORAB) have been reported to be 0.889 and 0.800, respectively, if the cutoff value for positive RPHA was defined as $2^1$ (Yoo, 1988). Correction with these values [number of both negatives (274,037) × negative predictability (0.8)] + [number of any positives (95,453) × {1-positive predictability (0.889)}] results in a decrease in the observation number of the denominator. In other words, if the RIA method including CORAB for anti-HBc had been introduced to sort out the pure negatives from HBV, the result might be greater than the results of the present study. It means that, in any case, the incidence rate of this study seems to underestimate the true value.

To observe the occurrence level of hepatitis B in a longitudinal manner, the study population must be restricted to the physically healthy person through periodic physical check-ups by a physician. All of them must be healthy at the time of recruitment, in terms of at least liver diseases. It is, therefore, certain that the incidence rate of this study must underestimate the true incidence of the whole population at the national level.

In most epidemiological studies on the characteristics of hepatitis B, the incident case was defined by seroconversion of HBV seromarkers (Beasley et al., West et al., 1986; Choi, 1986; Milne et al., 1987; Park et al., 1987). Moreover, reports on the occurrence of hepatitis B are extremely rare. Beasley et al. (1983) reported that among 738 college students susceptible to HBV in Taiwan, 5.3% had undergone serologic conversion, 12.8% of which were associated with hepatitis B for 3.5 years. The result can be converted into an annual incidence of 19.4/10,000, approximately 10 times higher than our results. Reasons for this seem to be due to the difference in age structure of population, to the definition of hepatitis B and to the real difference in incidence. However, it might be largely due to the total number of cases under observation, to the different selection criteria of the study population, and to the detection methods used for HBV seromarkers.

## Acknowledgements

Appreciation is expressed for the cooperation given by the Korea Medical Insurance Corporation (KMIC), and also for the technical assistance of the Systems Engineering Research Institute in the Korea Advanced Institute of Science and Technology (KAIST). The

authors would like to express their thanks to Mr. Poong-Min Kim and Mrs. Mi-Hyung Lee for their sincere efforts in data linking and data processing, as well as to the 17 senior medical students of the Seoul National University College of Medicine.

## REFERENCES

Ahn YO, Yoo KY: *An epidemiologic study on risk factors of HBs antigenemia in Korea. Korean Publ Health Assoc 9:207-217, 1983.*

Ahn YO, Yoo KY, Park BJ: *An epidemiologic study on risk factors of hepatitis B infection in Korea. Korean J Epidemiol 9:57-65, 1987.*

Beasley RP, Hwang LY, Lin CC, et al.: *Hepatocellular carcinoma and hepatitis B virus: a prospective study of 22,707 men in Taiwan. Lancet 2:1129-1132, 1981.*

Beasley RP, Hwang LY, Lin CC, et al.: *Incidence of Hepatitis B virus infection in preschool children in Taiwan. J Inf Dis 146:198-204, 1982.*

Beasley RP, Hwang Ly, Lin CC, et al.: *Incidence of hepatitis among students at a university in Taiwan. Amer J Epidemiol 117:213-222, 1983.*

Blumberg BS, Larouze B, London WT, et al.: *The relation of infection with the hepatitis B agent to primary hepatic carcinoma. Amer J Pathol 81:669-682, 1975.*

Breslow NE, Day NE: *Statistical methods in cancer research. Vol 1. The analysis of case-control studies. (IARC scientific publication number 32): Lyon: International Agency for Research on Cancer, 1980.*

Choi BY: *Incidence of hepatitis B virus infection of the schoolchildren in a rural area of Korea. Korean J Prev Med 19:281-292, 1986.*

Holford TR: *The analysis of rates and of survivorship using log-linear models. Biometrics 36:229-305 1980.*

Holford TR: *Strategies for the analysis of case-referent and cohort studies. In perinatal epidemiology edited by Bracken MB. New York, Oxford, Oxford University Press, 1984.*

Kim CY: *Prevention of viral hepatitis, type B, by vaccination with purified hepatitis B surface antigen. J Korean Med Assoc 22:1013-1025, 1979.*

Kim CY, Lee HS, Yu PC, et al.: *Comparison of radioimmunoassay and reversed passive hemagglutination assay in the detection of hepatitis B surface antigen. J Korean Med Assoc 27:49-53, 1984.*

Milne A, Allood GK, Moyes CD, et al.: *A seroepidemiological study of the prevalence of hepatitis B infection in a hyperendemic New Zealand community. Int J Epidemiol 16:84-90, 1987.*

Park BJ: *A study on the validity of passive hemagglutination test for HBsAb. Korean J Prev Med 20:114-119, 1987.*

Park JH, Youn TH, Chun BY, et al.: *Hepatitis B virus infection rate of medical school student in Taegu. Korean J Prev med 10:129-136, 1987.*

SAS Institute, Inc: *SAS/STAT guide for personal computers, Version 6 ed., Cary, NC: SAS Institute, Inc, 1987.*

Szmuness W: *Hepatocelluar carcinoma and hepatitis B virus: evidence for causal association. Prog Med Virol 24:40-69, 1978.*

The GLIM Working Party: *The generalized linear interactive modeling system, release 3.77. Oxford:, Numerical Algorithms Group, Ltd, 1987.*

West SK, Linago AL, Domingo Eo, et al.: *The Liver Study Group. Incidence and prevalence of hepatitis B. A community-based survey in the Philippines. Am J Epidemiol 123:681-689, 1986.*

Yoo KY: *Predictability of the test methods used in the detection for hepatitis B virus infection in a defined population. Seoul J Med 29:143-148, 1988.*

Yoo KY, Park BJ, Ahn YO: *Seroepidemiology of hepatitis B virus infection in healthy Korean adults in Seoul. Korean J Prev Med 21:89-98, 1988.*