BKM Public Health & Community Medicine
(Berita Kedokteran Masyarakat)

Volume 38 **Number** 4, 2022
Pages 107-112
DOI: 10.22146/bkm.v38i04.4413

Google Trends data analysis of hepatitis B and their correlation with hepatitis B early detection program for pregnant women in Indonesia

Dwi Adhi Nugraha[1], Triyanta Yuli Pramana[2]

Submitted:
April 13th, 2022
Accepted:
April 28th, 2022
Published:
April 30th, 2022

[1]RSU Fastabiq Sehat PKU Muhammadiyah Hospital, Pati, Indonesia

[2]Internal Medicine, Faculty of Medicine Sebelas Maret University, Central Java, Indonesia

***Correspondence:**
Dwi Adhi Nugraha
d.adhinugraha@gmail.com

Abstract

Purpose: Hepatitis B is still a problem in Indonesia. National surveillance project [Basic health survey (Riskesdas)] data in 2013 showed the prevalence of hepatitis B surface antigen (HBsAg) was 7.1%. One effort has been made to reduce the transmission of hepatitis B: the hepatitis B early detection program (DDHB) for pregnant women. Early changes in population health may be reflected in shifts in information and communication patterns on the internet. Changes in information and communication patterns on the internet can be an early "symptom" of changes in population health. This study aimed to know the correlation of hepatitis B relative search volume with HBsAg reactive percentage and coverage of the DDHB program for pregnant women in Indonesia. **Methods:** The value of relative search volume for hepatitis B in various regions in Indonesia in 2020 based on Google trends correlated with reactive HBsAg percentage and coverage of the DDHB program for pregnant women in 2020. Researchers compared the value of relative search volume for hepatitis B with hepatitis, hepatitis A, and hepatitis C. Spearman correlation analysis test was used to find the relationship between variables. **Results:** East Nusa Tenggara had the highest hepatitis B relative search volume and reactive HBsAg percentage in the DDHB program for pregnant women in 2020. Hepatitis B relative search volume was correlated with reactive HBsAg percentage ($r=0.618$; $p<0.001$) but did not correlate with DDHB coverage ($r=-0.65$; $p=0.374$). The popularity of hepatitis B (19.81±6.68) was higher than hepatitis A (8.65±5.79) and hepatitis C (3.90±1.75) but lower than hepatitis alone (54.08±19.02). **Conclusion:** East Nusa Tenggara had the highest hepatitis B relative search volume and reactive HBsAg percentage. Hepatitis B's relative search volume in Google Trends data analysis significantly correlated with reactive HBsAg percentage in the DDHB program for pregnant women in Indonesia.

Keywords: DDHB; Google Trend; hepatitis B; Indonesia

INTRODUCTION

Hepatitis B is still a significant health problem. There are 257 million people in the world with chronic hepatitis B infection. Twelve to 20% of patients will develop cirrhosis. Of these, about 20-23% develop decompensated cirrhosis, and 6-15% develop hepatocellular cancer [1]. Hepatocellular cancer is the fifth most common and second in terms of mortality [2]. In 2015, the World Health Organization (WHO) Southeast Asia Region (SEAR) reported that 40 million people suffered from chronic hepatitis B and 285,000 deaths related to complications of chronic infection, cirrhosis, and hepatocellular cancer [3]. Data in Indonesia compiled by Riskesdas 2013 from 33 provinces shows the prevalence of HBsAg is 7.1% (taken from 40,791 samples) [2].

Most hepatitis B transmission occurs in infancy during perinatal and early childhood transmission. To prevent this, WHO recommends that all infants receive at least three doses of the hepatitis B vaccine, including the birth dose [3]. In Indonesia, there are efforts to prevent transmission from women suffering from hepatitis B to their babies by carrying out early detection of hepatitis B (DDHB). A newborn child from a hepatitis B mother will be given hepatitis B immunoglobulin (HBIG) in addition to the hepatitis B vaccine [4].

Data related to hepatitis B, especially in the Asia Pacific region, is still insufficient to develop the proper intervention. In Indonesia, an inadequate disease surveillance system, geographic limitations for collecting data from more than 17,000 islands, and limited facilities for detecting chronic hepatitis B infection make hepatitis B underdiagnosed [2].

In recent years, internet data has been widely used to analyze and predict human behavior, including in the health sector [5]. Changes in information and communication patterns on the internet can be an early "symptom" of changes in population health. Furthermore, the difference between the prevalence of information and the disease incidence can be a marker of an information deficit [6]. Google Trends is the most widely used tool for dealing with health problems and topics using internet data [5]. Internet use in Indonesia is also increasing. According to the Indonesian Internet Service Providers Association (APJII), 171.7 million people, or around 64.8% of the total population of Indonesia, were using the internet in 2019. There was an increase of 10.2% compared to 2018 data [7].

Eyesenbach introduced the terms infodemiology and infoveillance to describe the use of data on the internet to identify and solve health problems. Infodemiology is “the science of distribution and determinants of information in electronic media, especially the internet, or within a population, to provide public health information and public policy” [6]. Infoveillance is “the activity of using infodemiology for surveillance purposes” [8].

Many information surveillance studies used Google trends data [5]. Google was Indonesia's most used search engine; about 97% of people used it. The majority of Internet users in Indonesia are adolescents [9]. Because of their risky behavior, adolescents are more vulnerable to hepatitis B infection [10]. Finding clues from this group may be sufficient to study hepatitis B information surveillance.

From the description above, we aim to analyze relative search volume data of hepatitis B internet searching in Indonesia using Google trends and then correlate it to the DDHB program data for pregnant women as an intervention to reduce hepatitis B transmission, especially vertical transmission.

METHODS

Google Trends is a website that provides information on popular topics based on Google users' searches. The data is openly available and accessible to everyone. Users can enter keywords and specify the desired region, period, and category. Google Trends is not case-sensitive but considers spelling errors that can occur. Google Trends will display numbers 0 – 100 based on the topic proportion for all search topics in a given region within a certain period. This scale shows a keyword's relative popularity based on the query's time and location. Different areas with the same search interest may not have the exact absolute search totals. A value of 0 does not mean any searches, but the search volume is too small to show results [5,9].

Researchers entered the keyword “hepatitis B” in 2020 in Indonesia and recorded the relative search volume values across all regions on a numerical scale. The study data included the top queries related to hepatitis B searches. Researchers also entered the keywords “hepatitis,” “hepatitis A,” and “hepatitis C” to compare with “hepatitis B” in 2020 in Indonesia.

Data related to the DDHB program for pregnant women were taken from the Indonesian Ministry of Health (http://sihepi.kemkes.go.id/dashboard_hepb). The percentage of reactive HBsAg and the percentage of pregnant women carrying out DDHB in all regions of Indonesia are recorded in numerical data.

The Shapiro-Wilk test assessed the normality of the hepatitis B relative search volume data, the reactive HBsAg percentage, and the DDHB coverage. The Pearson correlation test determines the relationship between variables if the data is usually distributed. If the data distribution is abnormal, the Spearman correlation test is performed. Statistical test using SPSS 23 program.

The comparison of relative search volume for hepatitis B, hepatitis, hepatitis A, and hepatitis C was done descriptively by calculating the average value. The top queries for the keyword hepatitis B were recorded in this study.

RESULTS

Table 1 shows that based on relative search volume values for the keyword hepatitis B on Google trends during 2020 from 31 regions in Indonesia, East Nusa Tenggara (NTT) had the highest volume (100), followed by Southeast Sulawesi (95), Central Sulawesi (65), Papua (63), and the Bangka Belitung Islands (62). The lowest relative search volumes were for Central Java (35), Riau (34), DKI Jakarta (33), West Java (32), and Gorontalo (31).

Table 1. Distribution of hepatitis B relative search volume data during 2020 based on Google trends, reactive HBsAg percentage, and coverage of the DDHB program among pregnant women in 2020

Province	Relative search volume value	HBsAg percentage (%)	DDHB coverage (%)
Aceh	47	1,01	65,98
Bali	37	1,23	56,05
Bangka Belitung Islands	62	2,15	70,93
Banten	36	1,53	52,2
Bengkulu	54	1,73	38,98
Central Java	35	1,28	72,21
Central Kalimantan	61	2,51	51,78
Central Sulawesi	65	2,48	50,42
East Java	40	1,87	68,6
East Kalimantan	61	1,92	55,93
East Nusa Tenggara	100	4,9	34,66
Gorontalo	31	3,27	56,49
Jakarta	33	1,39	58,98
Jambi	36	1,25	48,62
Lampung	39	1,35	66,37
Maluku	55	1,85	39,87
North Kalimantan	41	2,07	76,49
North Maluku	no data	3,9	48,67
North Sulawesi	58	1,54	18,68
North Sumatera	38	1,1	9,8
Papua	63	3,91	24,77
Riau	34	1,51	26,36
Riau Islands	50	1,55	38,38
Southeast Sulawesi	95	3,46	29,95
South Kalimantan	61	2,14	54,66
South Sulawesi	51	2,13	68,15
South Sumatera	38	0,72	41,32
West Java	32	1,25	49,24
West Kalimantan	53	2,41	41,76
West Nusa Tenggara	48	3	59,08
West Papua	no data	4,68	17,51
West Sulawesi	no data	2,2	47,92
West Sumatera	38	1,22	44,68
Yogyakarta	40	0,69	44,45
Indonesia	-	1,68	51,37

The top queries in the search for hepatitis B include hepatitis B disease (100), hepatitis B (71), hepatitis B vaccine (52), Hepatitis B drugs (43), and hepatitis B (40). Relative search volume for hepatitis B was higher than for hepatitis A and C but not higher than for hepatitis. The average value of hepatitis relative search volume was 19.81 ± 6.68, compared to hepatitis (54.08 ± 19.02), hepatitis A (8.65 ± 5.79), and hepatitis C (3.90 ± 1.75). The comparison is figured in Figure 1.

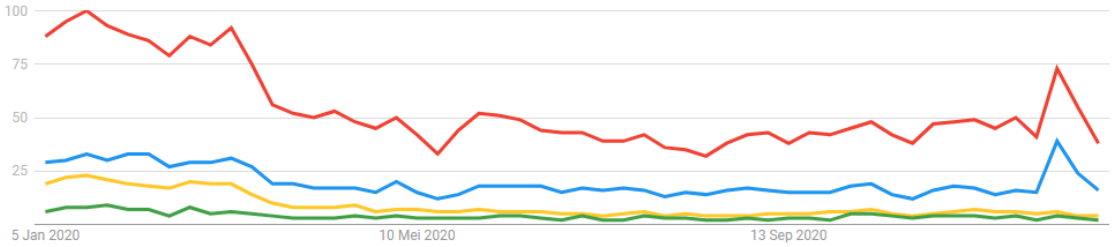

Figure 1. Comparison of relative search volumes for hepatitis B (blue), hepatitis (red), hepatitis A (yellow), and hepatitis C (green) during 2020

East Nusa Tenggara had the highest HBsAg prevalence, with 4.9%. The national average value of the HBsAg percentage was 1.68%. The highest rate of pregnant women who did the DDHB program was in North Kalimantan (76.49%). The national average of 51.37% was still lower than the desired target of 80%.

The results of the Shapiro Wilk test for hepatitis B relative search volume, reactive HBsAg percentage, and the coverage DDHB percentage were 0.06, 0.04, and 0.70, respectively. Because some data wasn't usually distributed, the researcher used the Spearman correlation test to find the relationship between variables.

Hepatitis B relative search volume had a statistically significant relationship with the percentage of reactive HBsAg (r=0.618; p<0.001). The relationship between hepatitis B relative search volume and DDHB coverage for pregnant women programs was not statistically significant (r= -0.65; p=0.374). The scatter plot graph between the relative search volume of hepatitis B and reactive HBsAg percentage is described in Figure 2.

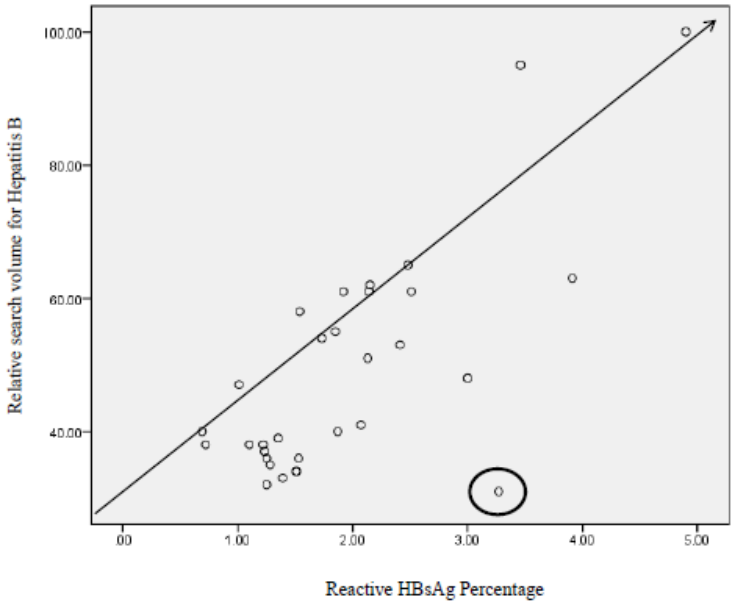

Figure 2. The scatter plot graph shows the relative search volume of hepatitis B and the reactive HBsAg percentage in DDHB for the pregnant women program in 2020. The circle shows the Gorontalo area with a reactive HBsAg percentage of 3.27 and a hepatitis B relative search volume value of 31.

DISCUSSION

More searches were conducted related to hepatitis B than hepatitis C in Indonesia. It could be due to the prevalence of hepatitis B being higher than hepatitis C. Riskesdas data in 2007 from 22 provinces showed the prevalence of reactive HBsAg was 9.4% (out of 10,391 samples), anti-HCV prevalence data was 0.82% (out of 11,762 samples). In 2013, based on RISKESDAS data, the prevalence of HBsAg from 33 provinces was 7.1% (out of 40,791 samples), while the prevalence of anti-HCV was 1% (out of 40,233 samples) [2].

East Nusa Tenggara had the highest hepatitis B relative search volume and reactive HBsAg percentage. East Nusa Tenggara is an archipelago province with a high transit area and can be easily reached from Timor Leste and Australia. The regions with rapid population movement increase hepatitis B transmission. Overall, HBsAg reactive incidence in east Indonesia is higher than in the west [10].

Hepatitis B attracted people more than hepatitis A and hepatitis C. This may be due to the greater public attention paid to hepatitis B. A semi-qualitative study in five Chinese cities that included 41 people with chronic hepatitis B infection showed that hepatitis B had a profound impact on personal and social life. There were changes in education choices, job choices, economic opportunities, and the development of intimate relationships. There is still stigma and discrimination against hepatitis B patients even though the government has implemented regulations to increase access to education and employment for hepatitis B patients [11].

Another public concern may arise from the pattern of hepatitis B transmission and its impact in the future. Transmission of hepatitis B can occur vertically from mother to child, especially in the perinatal period or early childhood, with a prognosis of becoming chronic, reaching 90%. Patients mostly remain asymptomatic and go undiagnosed thereafter [12]. Transmission of hepatitis B can also occur through sexual transmission [13]. If the infection becomes chronic, then 20% of patients can develop liver cirrhosis, and 6-15% of them progress to hepatocellular cancer [2].

We found that the relative search volume of hepatitis B significantly correlated with the percentage of reactive HBsAg in the DDHB program for pregnant women ($r=0.618$; $p<0.001$). Zhang et al. (2018) in China compared a model based on web search results to relate the actual incidence of hepatitis B sufferers. Based on the search volume on the Baidu Index from January 2011 to December 2016, the model was associated with the incidence of hepatitis B from September to December 2016 and the number of hepatitis B patients in January 2017. The relative error value of data comparison from the model and the actual number of hepatitis B was 0,42% - 1.46%. The adjusted R square model value of 0.891 ($p=0.001$) indicates that the model made from web searches is closely related to the increase in hepatitis B incidences, and the model can also be used for prediction [14]. Analyzing the pattern of internet searches related to hepatitis B through internet data can help us determine the increase in hepatitis B incidences in certain areas. Data analysis on the internet can be a solution to discover a disease's emergence or its prevalence quickly. It's more available in real-time than traditional data collection [5].

The increase in the number of web searches occurs when there is an increase in the number of diagnoses of certain diseases because the internet is the primary tool for people to seek health information. A qualitative study by Chu et al., conducted from November 2015 to January 2016 in Hong Kong on 49 respondents, was conducted to determine the public's perception of searching for health information online. All respondents agreed that they would use the internet for themselves or find information for others based on a desire to understand more, clarify, and confirm health problems. All respondents also agreed that they felt more comfortable looking for health information on the internet. Searching online is faster and can be done anywhere and anytime than accessing health services in conventional health facilities [15].

The scatter plot graph between the relative search volume and the HBsAg reactive percentage can show areas with high incidences but lack information, such as Gorontalo. It can warn health workers and public policymakers to make interventions to address the information gap. Using several metrics displayed on infodemiology dashboards can help public policymakers determine which areas require health campaigns [6].

Hepatitis B relative search volume had not significantly correlated with maternal participation in the DDHB programs ($r= -0.65$; $p=0.374$). Based on Permenkes number 53 of 2015 concerning the prevention of viral hepatitis, the DDHB program is offered to every pregnant woman who comes to a health facility for an integrated examination of hepatitis B and HIV [16]. The coverage target percentage is 80%, while in 2020, the national average coverage percentage was only 51.37%. To increase coverage of the DDHB program, it is necessary to provide adequate information about what the mother should do to prevent hepatitis B, especially transmission during the perinatal period. The attitude of pregnant women to participate in the DDHB program is influenced by their knowledge of hepatitis B itself. Providing information related to hepatitis B will significantly increase their knowledge and attitudes towards hepatitis B [16,17].

Of the top five queries related to hepatitis B, only one query related to hepatitis B prevention is the hepatitis B vaccine. Three were queries about hepatitis B in general. There was hepatitis B disease, hepatitis B is, and what is hepatitis B, and one other query related to hepatitis B treatment is the hepatitis B drugs.

Roughly, the public's search for efforts to prevent hepatitis B in pregnant women, such as HBsAg screening through the DDHB program, has not become a significant concern for the community. It is necessary to educate the public so that attention to the DDHB program increases and the percentage of mothers doing DDHB can reach the target.

This study has limitations, including that the search volume for the keyword "hepatitis B" was less than for "hepatitis," which can cause bias. Some areas did not have a high search popularity value in Google Trends, and access to the internet was not attributed to Indonesia. The relative volume of internet searches for hepatitis B based on Google trends correlated with the percentage of HBsAg in the DDHB program in Indonesia but not the percentage of DDHB coverage. Hepatitis B is still a significant public concern compared to other hepatitis.

The increasing use of the Internet encourages people to seek health information through it. Search volume patterns on the Internet can describe the incidence of disease. Discrepancies between the amount of information and disease incidence can indicate a lack of information in the community. They can help policymakers assess the health information the community needs to carry out the right interventions.

Google trends can be used to determine the incidence of a disease quickly. This can accelerate policymakers' intervention before data is collected, which can be advantageous considering Indonesia's geographical profile, which consists of many islands.

REFERENCES

1. Huang DQ, Lim SG. Hepatitis B: Who to treat? A critical review of international guidelines. Liver Int. 2020;40(S1):5–14.
2. H Muljono D. Epidemiology of Hepatitis B and C in Republic of Indonesia. Euroasian J Hepatogastroenterol. 2017;7(1):55-59
3. Sandhu HS, Roesel S, Sharifuzzaman M, Chunsuttiwat S, Tohme RA. Progress Toward Hepatitis B Control - South-East Asia Region, 2016–2019. MMWR Morb Mortal Wkly Rep. 2020;69(30):988–92.
4. Kementerian Kesehatan RI. Pencapaian DDHB [Internet]. Sihepi log-in. 2021 [cited 2022Mar29]. Available from: http://sihepi.kemkes.go.id/dashboard_hepb
5. Mavragani A, Ochoa G. Google Trends in Infodemiology and Infoveillance: Methodology Framework. JMIR Public Health Surveill. 2019;5(2):e13439.
6. Eysenbach G. Infodemiology and Infoveillance: Framework for an Emerging Set of Public Health Informatics Methods to Analyze Search, Communication and Publication Behavior on the Internet. J Med Internet Res. 2009;11(1):e11.
7. Wahyudiyono W. Implikasi Penggunaan Internet terhadap Partisipasi Sosial di Jawa Timur. J Komunika J Komun Media Dan Inform. 2019;8(2):63.
8. Eysenbach G. Infodemiology and Infoveillance. Am J Prev Med. 2011;40(5):S154–8.
9. Syamsuddin M, Fakhruddin M, Sahetapy-Engel JTM, Soewono E. Causality Analysis of Google Trends and Dengue Incidence in Bandung, Indonesia With Linkage of Digital Data Modeling: Longitudinal Observational Study. J Med Internet Res. 2020;22(7):e17633.
10. Kambuno NT, Waangsir F, Octrisdey K, Tangkelangi M, Ibraar M, Elengoe A. Serology Profile of Hepatitis B in Adolescents in Kupang, East Nusa Tenggara, Indonesia. 2021;5.
11. Wallace J, Pitts M, Liu C, Lin V, Hajarizadeh B, Richmond J, et al. More than a virus: a qualitative study of the social implications of hepatitis B infection in China. Int J Equity Health. 2017;16(1):137.
12. Umar M, Hamama-tul-Bushra, Umar S, Khan HA. HBV Perinatal Transmission. Int J Hepatol. 2013;2013:1–7.
13. Yang S, Wang D, Zhang Y, Yu C, Ren J, Xu K, et al. Transmission of Hepatitis B and C Virus Infection Through Body Piercing: A Systematic Review and Meta-Analysis. Medicine (Baltimore). 2015;94(47):e1893.
14. Zhang Y, Zhao Y, Lin X, Wang A, Che D. Modeling for the prediction of Hepatitis B incidence based on integrated online search indexes. Inform Med Unlocked. 2018;10:143–8.
15. Chu JT, Wang MP, Shen C, Viswanath K, Lam TH, Chan SSC. How, When and Why People Seek Health Information Online: Qualitative Study in Hong Kong. Interact J Med Res. 2017;6(2):e24.
16. Islahiyah N, Kholisotin K, Agustin YD. Pengaruh Paket Edukasi HbsAg terhadap Pengetahuan dan Sikap Ibu Hamil Trimester 1 tentang HbsAg. Indones J Health Sci. 2020;12(1):65–75.
17. Yanti NNL, Mahayati NMD, Armini NW. Pengaruh Penyuluhan dengan Media Video melalui Whatsapp Group tentang Hepatitis B terhadap Peningkatan Skor Pengetahuan Ibu Hamil. J Kebidanan Malahayati. 2021;7(4):824–9.