MJoSHT
Malaysian Journal of Science, Health & Technology
[mjosht.usim.edu.my]

Article

Advancing Localized Public Health Surveillance in Malaysia by Enhancing EIOS with Google COVID-19 Data Integration

Hazeeqah A.K. Aryffin[1], Sakinah Ali Pitchay[1,2], Murtadha A.B. Sahbudin[3], A.H. Azni[1,2] and Ilfita K. Sahbudin[4]

[1]*Faculty of Science and Technology, Universiti Sains Islam Malaysia, 71800, Nilai, Negeri Sembilan, Malaysia.*

[2]*CyberSecurity and Systems (CSS) Research Unit, Universiti Sains Islam Malaysia.*

[3]*Institute of Applied Data Analytics (IADA), Universiti Brunei Darussalam, Brunei.*

[4]*Rheumatology Research Group, Institute of Inflammation and Ageing, University of Birmingham, United Kingdom.*

Correspondence should be addressed to:
Sakinah Ali Pitchay; sakinah.ali@usim.edu.my

Article Info
Article history:
Received:15 November 2024
Accepted: 26 January 2025
Published: 18 February 2025

Academic Editor:
Muhammad Safwan Ibrahim

Malaysian Journal of Science, Health & Technology

MJoSHT2025, Volume 11, Issue No. 1
eISSN: 2601-0003

https://doi.org/10.33102/mjosht.v11i1.455

Abstract*— Epidemic intelligence has evolved from traditional manual reporting and field investigation methods to dynamic, real-time surveillance, driven by the 21st-century surge in digital data sources. Infectious diseases pose a significant global health threat, with traditional surveillance methods often facing delays in detecting and responding due to reliance on structured clinical data. The Covid-19 pandemic has emphasized the need for precise and actionable data to inform public health decisions. The current categorization of the Epidemic Intelligence from Open Sources (EIOS) system by country limits its ability to precisely track and monitor infectious diseases at more localized levels. This study focuses on enhancing the EIOS system by improving geographic data specifically for Malaysia. Currently, the EIOS platform, which incorporates data sources from Johns Hopkins University (JHU), the World Health Organization (WHO), and the Worldometers (WOM), provides country-level data that limits the effectiveness of localized interventions. This enhancement involves integrating state-level data in Malaysia from the Google Covid-19 Open Data Repository, which collects data automatically from authoritative sources, volunteers and contributors into the EIOS system. This paper presents the focus in Malaysia for better precision and effectiveness, in public health interventions. The suggested EIOS system will assist in sorting data based on dates,cases numbers,fatalities and data origins resultng in a more intricate and adaptable depiction of the pandemics advancement. The results of this study will offer insights and improvements for public health experts in Malaysia regarding the management and containment of infectious diseases at a local level.It will help optimize resource distribution and readiness efforts to mitigate the effects of outbreaks effectively.This improvement is intended to support targeted lockdowns and other public health interventions, in geographic areas with greater precision by enhancing geographical tracking capabilities.

***Keywords*— Epidemic Intelligence; COVID-19; Infectious Disease; Targeted Health Intervention; EIOS**

I. INTRODUCTION

Epidemic intelligence plays a role in monitoring public health to prevent the transmission of infectious diseases effectively as illustrated in Figure 1 depicting the World Health Organizations (WHO) strategy for tracking COVID 19 meticulously and systematically. This method allows for identification of affected areas, for segregation and testing to implement control measures and contain the spread efficiently by detecting cases within communities and healthcare facilities while maintaining structured data collection practices to track close contacts and monitor infected individuals closely. This method involves using sentinel surveillance to track disease trends in areas, with high risks.

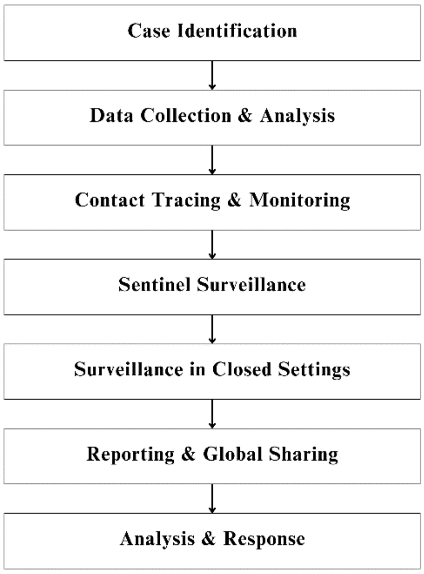

Figure 1. Sequential Framework of WHO-Recommended Public Health Surveillance for COVID-19 (WHO, 2022)

Public health monitoring in the past relied on information from facilities and research institutions alongside on site inquiries. Nowadays，gathering data from media platforms，cell phones， and digital health records has become a key method, for obtaining public health information (Agbehadji et al., 2020；Shausan et al., 2023).

In todays interconnected world where people travel frequently and goods are exchanged globally through means like direct interaction between individuals contaminated food and water airborne particles and carriers infectious diseases can easily spread from one person to another Public health monitoring and response protocols play a vital role in combating this global challenge As mentioned by Mboussou et al 2019 and Morgan & Pebody 2022 previous approaches often faced delays in identifying and addressing infectious diseases due, to their reliance solely structured clinical dataThus the capacity to thoroughly assess preventative measures, like social distancing and isolation is greatly restricted (Agbehadji et al., 2020; Ganser et al., 2022; Shausan et al., 2023).

The framework known as Epidemic Intelligence from Open Sources (EIOS) concentrates on diseases due to their significant health implications on a large scale. It underscores the nature of having efficient surveillance and response mechanisms in place (Lencucha & Bandara, 2021; Morgan & Pebody, 2022). EIOS sifts through open-source data streams to offer detailed insights while also working towards addressing the deficiencies found in current systems (MacIntyre et al., 2023; Shausan et al., 2023). The EIOS project was created in collaboration between the World Health Organization (WHO) and the European Commission’s Joint Research Centre (JRC) back in 2017. It has played a role, in enhancing worldwide health security by monitoring and examining publicly available data to identify and address potential public health risks (Lencucha & Bandara, 2021; Morgan & Pebody, 2022).

Twelve different networks partner with EIOS to manage data related to disease outbreaks effectively by collaborating with international health organizations and government bodies for timely responses to potential infectious diseases through the monitoring of news sources and social media alongside formal health reports using Artificial Intelligence (AI) (Bhatia et al., 2021; MacIntyre et al., 2023). AI plays a role in improving outbreak detection by utilizing machine learning techniques to analyze health data more efficiently than traditional case reports, in pinpointing unusual patterns. The ability to detect threats early is crucial, in preventing the spread of infections to become widespread epidemics or pandemics (Agbehaji et al., 2020).

The reliability of available data can vary. This poses challenges for EIOS in maintaining data accuracy (Spagnolo et al., 2020). Additionally, the system focuses primarily on a level when filtering information. This limitation hinders its capacity to offer insights tailored to individual cities or regions – crucial, for effectively addressing localized outbreaks (Lencucha & Bandara 2021).

The goal of this study is to improve EIOSs ability to detect and respond by incorporating information, from the states.

II. RELATED WORK

Investigative work is crucial, in handling public health crises. Event Based Surveillance (EBS) gathers data on incidents that could potentially endanger health helping to spot unusual health risks early on and stop them from escalating into widespread outbreaks (Kasamatsu et al., 2021; Yanagawa et al., 2022). AI examines huge amounts of unstructured data to identify early warning signs. Other notable epidemic intelligence systems are HealthMap and Early Warning and Response Surveillance (EWARS).

A. Overview of The EIOS System

The EIOS software processes millions of text entries from various sources in several languages. It prioritizes and classifies reports on health events, natural disasters, conflicts, and large gatherings by employing natural language processing (NLP), sorting, and labeling. These reports are reviewed by human analysts to ensure their accuracy and relevance before being shared with users (MacIntyre et al., 2023). EIOS supplements current official reporting channels (such as those mandated by International Health Regulations) by combining traditional monitoring with publicly available data. Rather than replacing existing systems, it aims to supplement human analysis and improve collaboration among global health specialists (Spagnolo et al., 2020).

B. Integration of EIOS with INFORM

EIOS is a component of the INFORM (Index for Risk Management) suite, which comprises risk analysis instruments such as the INFORM Risk Index to improves EIOS's risk assessment and response time to public health issues (MacIntyre et al., 2023; Spagnolo et al., 2020).

C. Integration of EIOS with INFORM

During the time of the COVID outbreak EIOS provided a public COVID dashboard that displayed data from trusted sources regarding COVID cases and fatalities (Ganser et al., 2022; MacIntyre et al., 2023; Yanagawa et al., 2022). This platform, with BlueDot played a crucial role in boosting surveillance efforts during significant occasions like the 2020 Olympic and Paralympics held in Japan (Kasamatsu et al., 2021; Shausan et al., 2023). The efficacy of these tools was showcased in improving monitoring in real time and safeguard public wellbeing, through the prompt and precise dissemination of information regarding a wide array of worldwide occurrences.

HealthMap is a system that was created in 2006 and works autonomously without intervention to keep track of various health occurrences such as non-contagious diseases using advanced filtering techniques and text analysis algorithms developed by Fisher and Robinson. It scans through 80 alerts related to infectious diseases every day and has been particularly recognized for its contribution, during the Ebola virus outbreak in the beginning of 2014 (MacIntyre et al., 2023).

The Centers for Disease Control and Prevention (CDC)'s Early Warning and Response Surveillance (EWARS) established in 2020 to identify uncommon incidents indicating possible outbreaks and facilitate quick interventions.The CDCs COVID - 19 Response Modeling Team leveraged EWARS data to determine mortality rates connected to COVID - 19 beyond mainland China by estimating the incubation period and duration until death while evaluating the likelihood of the virus spreading to the United States and other nations. This emphasizes the importance of having efficient and reliable systems, in place during outbreaks or health crises to aid in making well informed decisions (Ricks et al., 2022).

In summary, these systems highlight the significance of integrating cutting edge technologies into epidemic surveillance. EIOS, HealthMap and EWARS collaborate to improve the capacity to detect monitor and respond efficiently to public health risks. Through combining automated data processing with AI driven analysis they provide resources, for timely intervention aiding in controlling the transmission of infectious diseases and mitigating the effects of global public health emergencies (Kamarul Aryffin et al., 2024).

III. METHODOLOGY

In this part of the study process starts with analyzing and matching data from two sources. The Google COVID-19 Open Data Repository and EIOS. To ensure accuracy and consistency of information before merging the data for further examination. The following step involves studying the trends of COVID-19 epidemiology in Malaysia by looking into cases and deaths as well, as total cases and deaths while also investigating the case fatality rate (CFR). Every measurement is analyzed using methods to fully grasp how the pandemic has affected things. The research also looks into errors made when calculating the Case Fatality Rate (CFR) that might influence its precision. This strategy is geared towards presenting an dependable depiction of the data and its broader significance.

A. Data Comparison and Alignmen

In data analysis, the types of open-source data used by EIOS are examined to access relevant open-source APIs, identify and categorize data sources, and assess their contributions to EIOS's effectiveness in real-time disease monitoring. Since EIOS currently categorizes data by country, this study seeks to enhance its capabilities by incorporating Malaysia's state-level dataset from the Google COVID-19 Repository, enabling more precise tracking and monitoring of infectious diseases. This data is critical for tracking the pandemic's progression, comparing data accuracy across sources, and analyzing regional differences in case reporting. This phase offers an in-depth assessment of how each data source contributes to EIOS and evaluates their effectiveness. Tables 1 and 2 present COVID-19 case reports for Malaysia from EIOS and the Google COVID-19 Data Repository on February 4, 2020, respectively.

TABLE I. MALAYSIA COVID-19 CASE REPORT BY EIOS

date	cases	deaths	source	New cases	New deaths	New fatality pct
2020-02-04	8	0	JHU	0	0	0

TABLE II. MALAYSIA COVID-19 CASE REPORT BY GOOGLE COVID-19 DATA REPOSTORY

date	cumulative_ confirmed	cumulative _deceased	new_ confirmed	new_ deceased
4/2/ 2020	10	0	2	0

Table 1 presents COVID-19 data from EIOS, sourced from Johns Hopkins University (JHU), showing 8 total cases with 0 deaths, and no reports of new cases, new deaths, or new fatality percentages. In contrast, Table 2 offers data from the Google COVID-19 Data Repository, indicating 10 total cases with 0 deaths, including 2 new cases and 0 new deaths. These tables underscore the discrepancies in COVID-19 case reporting between the two sources. The differences in the number of cases reported on platforms show how challenging and inconsistent real time data reporting can be and stress the need to verify information from various sources to ensure accuracy. Both EIOS and the Google COVID-19 Data Repository are crucial for offering up to date data, on tracking COVID-19 spread and shaping public health plans. These variations highlight the intricacies of collecting and reporting data during a health crisis where individual sources may follow varying reporting criteria or schedules. To understand the reliability of the data and improve health responses effectively during global health emergencies like COVID 10 crisis in Malaysias Johor region on February 4th, 2020 is essential to analyze these data variations, for consistency and actionability to harmonize and standardize data.

TABLE III. GOOGLE COVID-19 REPOSITORY MY (STATE: JOHOR) BEFORE DATA MAPPING

Location key	date	Cumulative confirmed	Cumulative deceased	New confirmed	New deceased
MY_01	4/2/ 2020	7	0	2	0

TABLE IV. GOOGLE COVID-19 REPOSITORY MY (STATE: JOHOR) AFTER DATA MAPPING

date	cases	deaths	source	New cases	New deaths	New fatality pct
2020-02-04	7	0	Google	0	0	0

Table 3 displays the information prior to mapping with columns labeled as "cumulative_confirmed", "cumulative_deceased", "new_confirmed", and "new_deceased". Following the data mapping process in Table 4 reveals the data but with standardized column headings as "cases", "deaths", "new_cases", and "new_deaths". These tables emphasize the transition, to consistent and clear column names post data mapping for improved reporting accuracy. The research delves into the exploration of particular algorithms and analytical techniques used by EIOS system including thorough data analysis and evaluation of algorithms and machine learning methods along with a comprehensive description of the strategies utilized by EIOS system. This inquiry seeks to grasp the methods that enhance epidemic intelligence within the EIOS framework. EIOS predominantly utilizes R, for computation and visual analysis to meet the data analysis needs for epidemiological monitoring purposes. Furthermore the project utilizes *Shell* scripts to tasks and enhance workflow efficiency; *Makefile* is employed for handling the build process. EIOS gathers its information from JHU WHO, the European Centre, for Disease Prevention and Control (ECDC) and Worldometers (WOM).

TABLE V. ALGORITHM 1 EXAMPLE OF DATA SANITIZATION IN WOM

Algorithm 1 Example of Data Sanitization in WOM

```
Start
    for each column i=2 to n do
        If data[i] = "" or data[i] = "-"
            data[i] = NA
        Endif
        remove all ',' from data[i]
        remove all '+' from data[i]
        // attempt to convert data to integer
        tmp = convert data[i] to integer
        //check for conversion warning
        If tmp has warning then
            data[i] = original format
            //return data in original format if conversion fails
        Else
            data[i] = tmp
            //update data in column i to integer
        Endif
    Endfor
End
```

In table 5, algorithm 1, data cleaning and validation for columns from the second onward are performed on data sourced from WOM before display. The process handles missing values and data conversion through a loop that iterates over each column. This loop replaces empty strings and "-" with NA, removes commas and plus signs, and converts the data to integers while managing any warnings that arise. Automated data-fetching scripts retrieve new data, which then undergoes the same validation, cleaning, and transformation procedures.

B. COVID-19 Epidemiological Trends in Malaysia

Malaysia reported its first COVID-19 case on January 25, 2020, with Seremban, the capital of Negeri Sembilan, recording its initial case on February 5, 2020. The National COVID-19 Immunization Programme commenced on February 24, 2021, with vaccinations beginning in Negeri Sembilan on March 3, 2021. The program provided free vaccines in three phases: Phase 1 targeted front-line healthcare workers, Phase 2 began on April 19, 2021, for elderly adults and high-risk groups, and Phase 3 opened on July 12, 2021, for all eligible individuals over 18 (Amin et al., 2023). An 8-week interval between the first and second doses of COVID-19 vaccines may be beneficial for some individuals, as it could reduce the small risk of myocarditis and pericarditis associated with these vaccines (CDC, 2024). The National COVID-19 Immunization Programme for Booster doses, or PICK-B, was launched on October 13, 2021, with the goal of providing booster doses to around 23 million fully vaccinated individuals in Malaysia (Hamdan et al., 2024).

The outbreak of COVID-19 in Malaysia caused changes in the healthcare and technology industries and raised considerable doubts among the population about what lay ahead.The heightened instability during March 2020 emphasized the role of up to date information, in dealing with unforeseen circumstances (Mahussin et al., 2021).

Malaysia’s Movement Control Order (MCO) saw a study revealing public awareness and compliance with COVID-19 protocols. However certain gaps in readiness were still evident. The public relied heavily on broadcast media, for COVID-19 information highlighting the vital role of digital platforms in educating the populace (Syed Mohamad et al., 2021).

The outbreak of COVID-19 brought changes to how consumers, in Malaysia approached online shopping. While people were informed, they may not have been entirely prepared to navigate the prolonged challenges of the pandemic. As individuals grew more cautious about in-person interactions, online platforms emerged as the preferred method for purchasing goods (Mohamad Shariff & Nur Hayani Izzati Abd Hamid, 2021). The pandemic require educational institutions in Malaysia to transition to online teaching and learning. Teachers encountered various challenges, such as technical barriers and increased workloads, but gradually adapted to the new online teaching methods (Ismail et al., 2021).

When people are well-informed about risks and necessary precautions, they are more likely to adjust their routines to reduce exposure, adapt to new norms, and comply with public health guidelines. Increased awareness also encourages greater use of digital tools for remote learning, working, and shopping, effectively reducing the need for physical interaction. Therefore, the capacity to raise and sustain awareness through accurate data and timely information is essential for guiding public behavior and mitigating the effects of a health crisis.

1) Analysis of New COVID-19 Cases: Figure 2 shows the trend in new COVID-19 cases in Malaysia from January 26, 2020, to June 15, 2024, revealing several major surges, each followed by rapid declines. These sharp decreases, particularly after peak periods, emphasize the positive impact of vaccination programs. The recurring pattern of case rises followed by substantial drops suggests that widespread vaccine administration played a crucial role in controlling the virus's spread following initial outbreaks. By October 2021, the reduction in new COVID-19 cases was largely due to the extensive administration of first and second vaccine doses. To prevent further surges, Malaysians were encouraged to get booster doses starting in October 2021. Booster uptake increased significantly towards late 2021 and early 2022, contributing to another reduction in case numbers by April 2022 as cases had started to climb again.

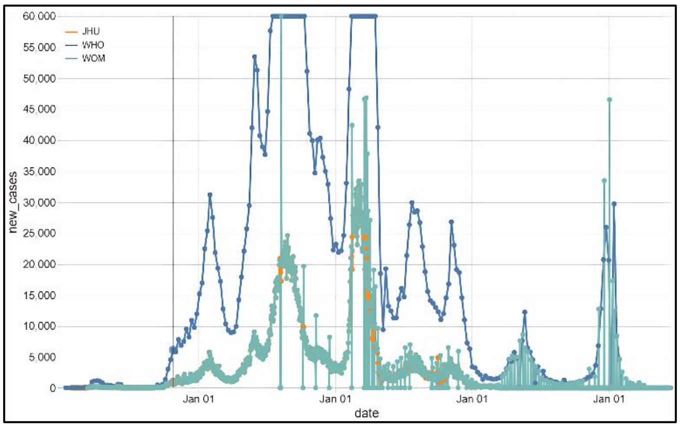

Figure 2. Number of New COVID-19 Cases in Malaysia from January 2020 until June 2024

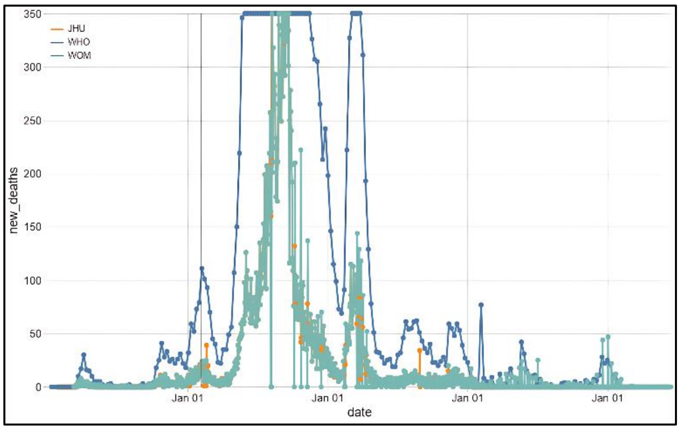

Figure 3. Number of New Deaths caused by COVID-19 in Malaysia from January 2020 until June 2024

TABLE VI. ALGORITHM 2 PSEUDOCODE FOR NEW CASES (ΔC)

Algorithm 2 Pseudocode for New Cases (ΔC)

```
Start
    for each day i=1 to n do
        If i=1 then
            new_cases[i] = cumulative_cases[i]
        Else
            new_cases[i] = cumulative_cases[i] -
            cumulative_cases[i-1]
        Endif
    Endfor
End
```

TABLE VII. ALGORITHM 3 PSEUDOCODE FOR NEW DEATHS (ΔD)

Algorithm 3 Pseudocode for New Deaths (ΔD)

```
Start
    for each day i=1 to n do
        If i=1 then
            new_deaths[i] = cumulative_ deaths[i]
        Else
            new_deaths[i] = cumulative_deaths[i] -
            cumulative_deaths[i-1]
        Endif
    Endfor
End
```

Table 6, algorithm 2 calculates the number of new COVID-19 cases based on cumulative case data. For each day, *i* from 1 to *n*, if *i* = 1, the number of new cases on that day is set to the cumulative cases on that day. For subsequent days, the number of new cases is determined by subtracting the cumulative cases of the previous day from the cumulative cases of the current day.

$$\Delta C(t) = C(t) - C(t-1) \tag{1}$$

This process is summarized by the equation above where $\Delta C(t)$ represents the number of new cases on day t, $C(t)$ is the cumulative number of cases on day t, and $C(t-1)$ is the cumulative number of cases on the previous day. In other words, subtracting today's cumulative number of cases from yesterday's cumulative number of cases will give the number of today's new cases. This equation and the pseudocode together provide a clear method for computing daily new cases from cumulative data.

2) Analysis of New COVID-19 Deaths: Figure 3 depicts the trend in new COVID-19-related deaths in Malaysia from January 26, 2020, to June 15, 2024. The graph shows several major peaks in new deaths, each followed by steep declines. Initially, as COVID-19 cases rose, the government launched vaccination programs to combat the virus. A significant reduction in new deaths began in November 2021, and, following the introduction of booster doses, a sharp decrease in fatalities was observed, especially by March 2022. This pattern highlights the effectiveness of vaccination and booster programs in reducing COVID-19-related mortality rates.

Table 7, algorithm 3 provides the pseudocode for calculating the number of new deaths (ΔD) due to COVID-19 based on cumulative death data. For each day, *i* from 1 to *n*, if *i* =1, the number of new deaths on that day is set to the cumulative deaths on that day. For subsequent days, the number of new deaths is determined by subtracting the cumulative deaths of the previous day from the cumulative deaths of the current day.

$$\Delta D(t) = D(t) - D(t-1) \tag{2}$$

This can be mathematically expressed with the equation above where ΔD represents the number of new deaths on day t, $D(t)$ represents the cumulative number of deaths on day t, and $D(t-1)$ represents the cumulative number of deaths on the previous day. This equation and the pseudocode together provide a clear method for computing daily new deaths from cumulative data.

3) Analysis of Cumulative COVID-19 Cases: Figure 4 shows the cumulative number of COVID-19 cases in Malaysia from January 26, 2020, to June 15, 2024. The graph reveals a rapid rise in cases during the early stages of the pandemic, especially during major outbreaks. However, from April 2022 onward, the rate of new cases began to slow, aligning with Malaysia's shift to the endemic phase on April 1, 2022 (Amira et al., 2023; Md Nadzri et al., 2024).

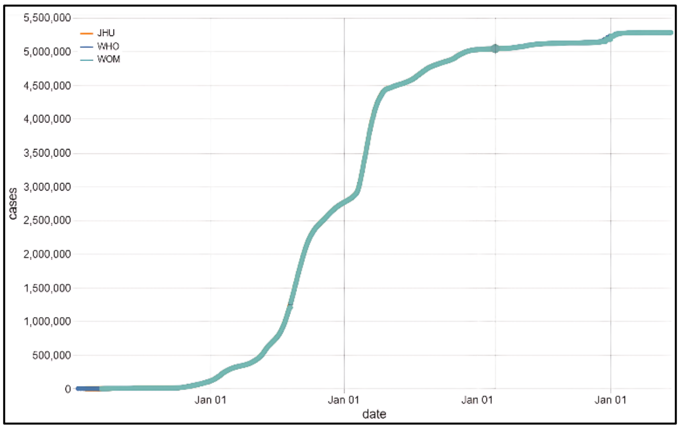

Figure 4. Cumulative Number of COVID-19 Cases in Malaysia from January 2020 until June 2024

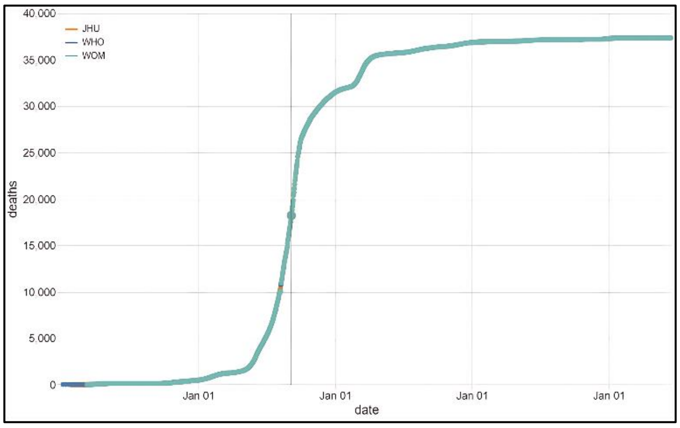

Figure 5. Cumulative Number Deaths caused by COVID-19 in Malaysia from January 2020 until June 2024

During this transition, numerous restrictions were eased as part of the government's strategy to adapt to living with the virus. For instance, face masks became optional for outdoor activities, though they remained mandatory in indoor settings and on public transport. Additionally, measures such as mandatory QR code scanning via the MySejahtera mobile app and vaccination status checks for entry into premises were discontinued (Amira et al., 2023).

TABLE VIII. ALGORITHM 4 PSEUDOCODE FOR CUMULATIVE CASES (C)

Algorithm 4 Pseudocode for Cumulative Cases (C)

```
Start
    for each day i=1 to n do
        If i=1 then
            cumulative_cases[i] = new_cases[i]
        Else
            cumulative_cases[i] =
            cumulative_cases[i-1] + new_cases[i]
        Endif
    Endfor
End
```

Table 8, algorithm 4 provides the pseudocode for calculating cumulative COVID-19 cases (*C*). For each day *i* from 1 to *n*, if *i*=1, the cumulative cases on that day, are set equal to the number of new cases for that day. For subsequent days, the cumulative cases are calculated by adding the new cases of the current day to the cumulative cases of the previous day.

$$C(t) = C(t-1) + \Delta C(t) \quad (3)$$

This algorithm is captured mathematically in the equation above where $C(t)$ represents the cumulative cases on day t, $C(t-1)$ represents the cumulative cases on the previous day, and $\Delta C(t)$ represents the new cases on day t. This process ensures that the cumulative total is updated daily, reflecting the continuous increase in cases over time.

4) Analysis of Cumulative COVID-19 Deaths: Figure 5 illustrates the cumulative number of deaths caused by COVID-19 in Malaysia from January 26, 2020 until June 15, 2024. The graph shows a steep increase in deaths during the initial waves of the pandemic, reflecting the high mortality rate during those periods. Despite the significant uptake of booster shots among the adult population, the number of deaths surged again in March 2022.

This increase in mortality was particularly marked among the elderly, especially those who had not received full vaccination and booster doses. Older individuals, especially those over 80, are more susceptible due to factors like weakened immunity, chronic health conditions, and overall frailty, which slow their recovery from COVID-19 and increase the chances of severe or fatal outcomes. After Malaysia transitioned to the phase on 1st April of this year there was a noticeable decrease in the number of deaths with it frequently staying below 10 deaths per day. This drop is mainly credited to the decline in new cases during the endemic period widespread compliance, with health recommendations endorsed by the Ministry of Health and the effects of the country's COVID-19 vaccination initiatives. The significance of this pattern highlights the necessity of vaccination and public health strategies in handling and diminishing the effects of the virus on, at risk populations (Amira et al., 2023).

TABLE IX. ALGORITHM 5 PSEUDOCODE FOR CUMULATIVE DEATHS (D)

Algorithm 5 Pseudocode for Cumulative Deaths (D)

```
Start
    for each day i=1 to n do
        If i=1 then
            cumulative_deaths[i] = new_deaths[i]
        Else
            cumulative_deaths[i] =
            cumulative_deaths[i-1] + new_deaths[i]
        Endif
    Endfor
End
```

Table 9, algorithm 5 outlines the step-by-step instructions for calculating the number of deaths (*D*) attributed to COVID-19 over time progression due to the ongoing pandemic situation. The initial step involves going through each day from day one up to a specific day denoted as *n*. When starting the calculation process at day one (denoted as *i=1*) the cumulative deaths recorded are based upon the number of new deaths reported for that day. Subsequently moving forward with each passing day in the series of calculations requires updating the cumulative death count by incorporating the number of new deaths reported for that specific day with the existing cumulative figure, from previous days. This approach ensures that an accurate record is maintained concerning the number of fatalities associated with COVID-19 while reflecting its continual impact throughout the pandemic period.

$$D(t) = D(t-1) + \Delta D(t) \quad (4)$$

The mathematical expression above showcases how the process is depicted in detail. In the equation provided; *D(t)* signifies the deaths on day *t*; *D(t−1)* signifies the total deaths on the previous day; and *ΔD(t)* represents the new deaths on day *t*. This formula captures the essence of the pseudocode by offering a straightforward way to compute cumulative deaths through the incremental addition of daily new death counts. This approach is vital for monitoring the overall death count over time—an imperative aspect, in grasping the gravity and evolution of the current pandemic situation.

5) Analysis of Case Fatality Rate (CFR): Figure 6 shows the COVID-19 CFR in Malaysia spanning from January 2020 to June 2024. A measure showing the percentage of deaths among confirmed cases during this period. The peak of CFR was observed in March and April 2020 at 1.7% reflecting the initial challenges faced during the early stages of the pandemic when healthcare systems were strained, and treatment options were limited. Subsequently from that point in October 2020 CFR decreased to below 1% attributable, to enhanced medical responses implemented alongside public health measures and successful vaccine distribution efforts.

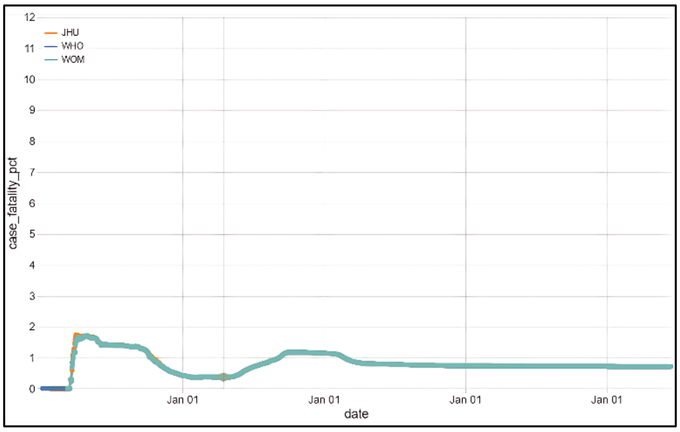

Figure 6. Case Fatality Rate of COVID-19 in Malaysia from January 2020 until June 2024

However, the CFR saw an increase again from September 2021 to February 2022, rising above 1.0%. This period coincided with the emergence of new COVID-19 variants and the lag in booster shot administration, which likely contributed to higher mortality among the infected population. Despite these fluctuations, the overall trend shows a stabilization of the CFR as the country continued to manage the pandemic through vaccination efforts and other public health strategies.

TABLE X. ALGORITHM 6 PSEUDOCODE FOR CUMULATIVE FATALITY RATE (CFR)

Algorithm 6 Pseudocode for Cumulative Fatality Rate (CFR)

```
Start
    for each day i=(L+1) to n do
        If cases[i-L] then
            CFR[i] = (deaths[i] / cases[i-L]) * 100
        Else
            CFR[i] = 0 // to handle division by zero by setting CFR to 0
        Endif
    Endfor
End
```

6) Recommended CFR Method: Table 10, algorithm 6 provides the pseudocode for calculating the CFR of COVID-19. The process begins by iterating through each day i from L+1 to n, where L represents the lag period, which is the average time from case confirmation to death. If the number of cases on day i−L is non-zero, the CFR for day i is calculated by dividing the number of deaths on day i by the number of cases on day i−L and then multiplying by 100 to express the result as a percentage. This ensures that the fatality rate is based on the appropriate time-lagged cases, reflecting a more accurate death-to-case ratio.

$$CFR = \frac{D(t)}{C(t-L)} * 100 \quad (5)$$

Worldometer proposed the CFR equation as shown above where *D(t)* represents the cumulative number of deaths on day *t*, and *C(t−L)* represents the cumulative number of cases on day $t-L$. This equation accounts for the time delay between when cases are confirmed and when deaths occur, providing a more accurate representation of the fatality rate over time. By adjusting for this lag, the CFR calculation becomes more reliable and reflective of the true risk of death associated with COVID-19 (Worldometer, 2024).

The flawed approach to calculating the case fatality rate involves simply dividing the number of deaths by the number of cases on the same day. For example, using this flawed method, the CFR is calculated as below.

$$CFR = \frac{D(t)}{C(t)} * 100 \quad (6)$$

However, this approach to calculating CFR is flawed because it assumes all deaths occur within the same timeframe as reported cases, which is rarely true for COVID-19 or other diseases. The time from infection to death varies and using the case count from the same day results in an inaccurate, often understated fatality rate. Since deaths are a lagging indicator, occurring days or weeks after infection, this method can create a misleading picture of the disease's severity. Sometimes dismissing this delay can lead to an estimation of the mortality rate of the virus and may inadvertently cause a sense of complacency in how public health measures are approached and the true extent of the virus's effects being reported accurately. For accuracy in calculation accuracy of CFR a more detailed approach involves taking into account the average duration from confirming a case to the occurrence of death denoted as " L". This method links deaths to the cases that probably led to them instead of combining deaths, with all cases reported within the same day.

The improved approach now captures the progression of the illness by taking into consideration the lag time between contracting the virus and succumbing to it—a more practical gauge of COVID-19's deadliness. This adjusted calculation can assist health officials in gauging the outbreaks seriousness and deploying resources effectively while putting in place tailored strategies to mitigate the virus's effects. Precise calculations of CFR are crucial, for comprehending and conveying the risks associated with COVID-19 to both the public and decision makers (Soliman Ashraf et al., 2020).

IV. RESULTS AND DISCUSSION

The EIOS system has been improved to analyze COVID-19 data at a detailed geographical level, in Malaysia for targeted public health efforts. This research incorporates information from all twelve states in Malaysia (Johor, Kedah, Kelantan, Kuala Lumpur, Labuan Melaka, Negeri Sembilan Pahang, Perak, Perlis, Pulau Pinang, Putrajaya, Sabah， Sarawak, Selangor, and Terengganu) into the EIOS framework; previously it mainly focused on data sources like JHU, WHO, and WOM.

By incorporating state-level information from the Google COVID-19 Open Data Repository into the system capabilities to monitor COVID-19 cases, with precision has enabled targeted healthcare responses. The upgrade enables interventions in areas at risk while keeping interruptions to a minimum in regions, with fewer cases. Users also have the option to sort data based on dates, cases and deaths and sources to personalize their understanding of how the pandemic's evolving.

Figure 7 depicts the COVID19 infections, in Johor sourced from the Google COVID19 Open Data Repository and showcased on the EIOS Dashboard. This visualization accentuates patterns in cases providing a snapshot of the pandemic's effects, in Johor. With this data, public health officials can monitor COVID-19 at the state level more precisely.

This case report highlights prolonged presentation of post-COVID-19 infection. Even though these symptoms are deemed mild, the patients are affected psychologically and functionally by this condition. Therefore, proper follow-up and individualized management are important approaches to this problem.

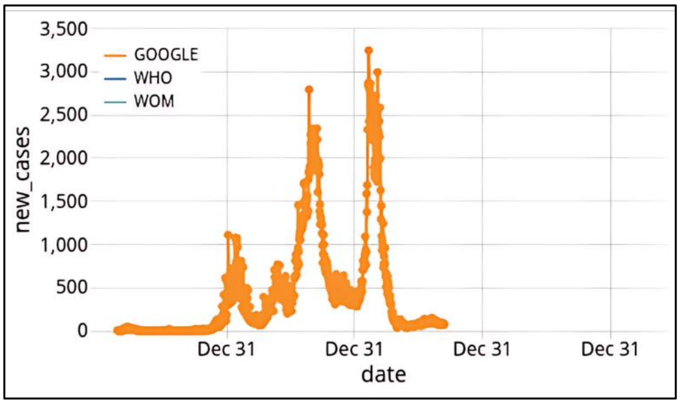

Figure 7. Number of New COVID-19 Cases in Johor, Malaysia from January 2020 until September 2022

The inclusion of Johor's data into the EIOS Dashboard through the Google COVID-19 Repository enhances the ability to monitor fluctuations in case numbers, identify peaks, and assess the effectiveness of public health interventions specific to Johor. This integration enables quicker responses, to outbreaks by allowing the EIOS system to provide customized insights based specifically for each states circumstance.

Next, the inclusion of cumulative COVID-19 cases from the Google COVID-19 Repository for Johor is illustrated in Figure 8. This figure complements previous visualizations by offering a cumulative perspective rather than just new case trends, further refining the insights available to public health officials. This localized focus on Johor not only highlights regional case trends but also underlines the importance of state-level data in facilitating more nuanced public health responses.

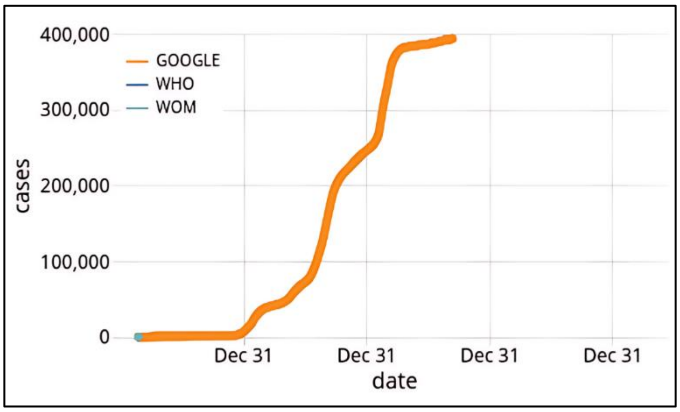

Figure 8. Number of Cumulative COVID-19 Cases in Johor, Malaysia from January 2020 until September 202

Additionally, Figure 9 displays daily COVID-19 deaths in Johor from the Google COVID-19 Repository. This daily tracking of number of deaths emphasizes the importance of epidemic monitoring to enable timely, informed decision-making to protect public health.

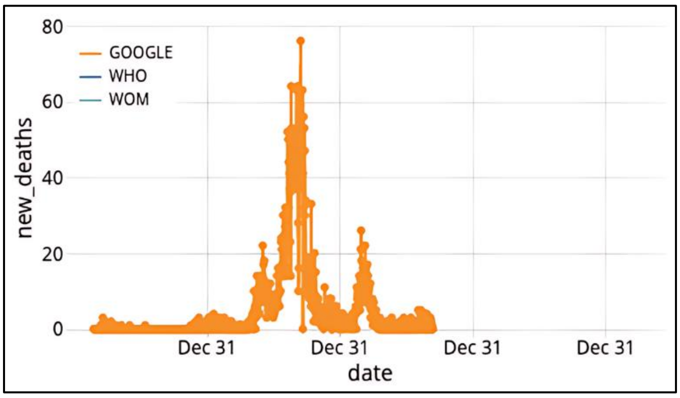

Figure 9. Number of New COVID-19 Deaths in Johor, Malaysia from January 2020 until September 2022

Lastly, Figure 10 presents the cumulative total of COVID-19 deaths in Johor, offering insight into the pandemic's long-term impact. This cumulative perspective highlights the lasting impact COVID-19 has taken on the state, underscoring the overall severity of the crisis. This long-term view complements the daily data by showing the broader, cumulative impact of the pandemic on the community.

While there may be minor differences in the data provided by sources like the Google COVID-19 Repository and the EIOS system, as shown in Tables 1 and 2, these variations are generally small and can be reconciled through careful analysis. The benefits of integrating localized data provide a more nuanced understanding of the pandemic's impact across different regions. By using information, from outlets and databases EIOS can provide a detailed and precise view promoting the overall success of public health strategies.

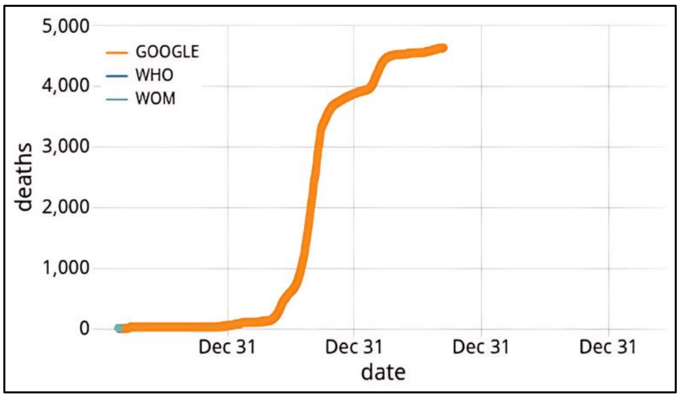

Figure 10 COVID related deaths, in Johor,Malaysia from January 2020 to September 2020

In conclusion, incorporating data, into EIOS will enhance Malaysia's capability to track and address COVID 19 at a level effectively ensuring that interventions are timely and tailored to each state's specific conditions. Additionally, health authorities can identify states that may require quarantine measures enabling a precise response. This tailored strategy aids in minimizing disruptions, in low case areas while allocating resources to the regions that require them the most. By improving the response, in this manner not only enhance the pandemic response but also establishes a stronger public health system that can effectively handle upcoming health crises.

V. FUTURE WORK AND CONCLUSIONS

Integrating state-level data from the Google COVID-19 Open Data Repository into EIOS has greatly improved Malaysia's ability to monitor and manage localized outbreaks of infectious diseases. By narrowing the geographic focus, the enhanced EIOS system gives public health officials the insights to implement targeted interventions and allocate resources more effectively.

One important feature for future work on the EIOS dashboard is the incorporation of city-level data into the system which can assist with better visualization of trends. For example, data observations from each city might improve the systemic capabilities for local demographic focusing. The addition of city-based trends data would increase the number of local trends that would be utilized by the EIOS dashboard, making it possible to use the data to better organize responses, monitoring and resource utilization aimed at specific areas within a city.

Next feature is to improve EIOS analytical functions. The system could be able to more accurately identify potential hotspots, trends, and at-risk populations by employing cutting-edge analytics and artificial intelligence. Such predictive features would be a great help to public health officials since these would provide actionable insights that should be acted upon to avert the escalation of outbreaks.

Furthermore, enhancing the EIOS dashboards could significantly improve user experience and satisfaction by better addressing the diverse needs of its users. Users could interact better with the data through customizable dashboards, alerts, and maps and thus enhance the public's engagement with the health system. In this way, reinforcing and broadening its usability and reach, the EIOS dashboard would be crucial in strengthening the health system in Malaysia for many years to come.

The EIOS framework's consistent data processing and visualization methods were crucial in providing accurate, up-to-date information. Such improvements facilitate targeted actions, enabling better use of resources – increasing efficiency of available resources in response to both existing and future health threats. Given the emerging public health threats, it is even more critical for EIOS to integrate granular, real-time data supported by AI for monitoring and outbreak prediction. This not only improves the capacity to manage health challenges presently, but also improves the public health system of Malaysia in all aspects this would make Malaysia stronger in dealing with the health threats that will arise. These developments provide the basis for EIOS to be a strong tool for safeguarding the health of the population and strengthening the country's health response in the future of Malaysia.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this paper.

ACKNOWLEDGEMENT

This research was funded by the Ministry of Higher Education (MOHE) Malaysia under the Fundamental Research Grant Scheme (FRGS/1/2023/ICT03/USIM/02/3). The authors would like to thank Universiti Sains Islam Malaysia (USIM) and MOHE for the support and facilities provided.

REFERENCES

[1] Agbehadji, I. E., Awuzie, B. O., Ngowi, A. B., & Millham, R. C. (2020). Review of big data analytics, artificial intelligence and nature-inspired computing models towards accurate detection of COVID-19 pandemic cases and contact tracing. In *International Journal of Environmental Research and Public Health* (Vol. 17, Issue 15, pp. 1–16). MDPI AG. https://doi.org/10.3390/ijerph17155330

[2] Alo, U. R., Nkwo, F. O., Nweke, H. F., Achi, I. I., & Okemiri, H. A. (2022). Non-pharmaceutical interventions against covid-19 pandemic: Review of contact tracing and social distancing technologies, protocols, apps, security and open research directions. In *Sensors* (Vol. 22, Issue 1). MDPI. https://doi.org/10.3390/s22010280

[3] Amin, K. H. A. K., Bujang, N. N. A., Abas, S. A., Zulkifli, N. I., Amir, S. A., Shah, S. M., Ganesan, V., Aziz, N. F., Jalaluddin, M. A., Wahil, M. S. A., Hasani, M. H. M., Ali, N. K. M., & Yusof, M. P. (2023). Epidemiology of COVID-19 cases and vaccination coverage in Seremban District, Malaysia, 2021. *Western Pacific Surveillance and Response Journal*, *14*(2). https://doi.org/10.5365/wpsar.2023.14.2.985

[4] Amira, N., Rosday, A., Azlina, K., Akil, K., Bidin, J., Sharif, N., & Chulan, M. (2023). Analyzing the Covid-19 Cases in Malaysia during the Transition to Endemic Phase. In *Applied Mathematics and Computational Intelligence* (Vol. 12, Issue 2).

[5] Bhatia, S., Lassmann, B., Cohn, E., Desai, A. N., Carrion, M., Kraemer, M. U. G., Herringer, M., Brownstein, J., Madoff, L., Cori, A., & Nouvellet, P. (2021). Using digital surveillance tools for near real-time mapping of the risk of infectious disease spread. *Npj Digital Medicine*, *4*(1). https://doi.org/10.1038/s41746-021-00442-3

[6] CDC. (2024). *Interim Clinical Considerations for Use of COVID-19 Vaccines in the United States*. CDC. https://www.cdc.gov/vaccines/covid-19/clinical-considerations/interim-considerations-us.html

[7] Ganser, I., Thiébaut, R., & Buckeridge, D. L. (2022). Global Variations in Event-Based Surveillance for Disease Outbreak Detection: Time

Series Analysis. *JMIR Public Health and Surveillance*, *8*(10). https://doi.org/10.2196/36211
[8] Hamdan, N. E. A., Yassen, A. O., & Loganathan Fahrni, M. (2024). COVID-19 Booster Vaccination in Malaysia. *Journal of Clinical and Health Sciences*, *9*(1), 6–14. https://doi.org/10.24191/jchs.v9i1.19570
[9] Ismail, S., Mohammad Ilham Husaini Mohammad Nazeri, & Sharifah Fairuz Syed Mohamad. (2021). The Impact of Covid-19 on Online Teaching and Learning (TnL) Towards Teachers in Malaysia. *Malaysian Journal of Science Health & Technology*, *7*(2), 15–22. https://doi.org/10.33102/mjosht.v7i2.169
[10] Kamarul Aryffin, H. A., Sahbudin, M. A. Bin, Ali Pitchay, S., Ab Halim, H., & Kamaliah Sahbudin, I. (2024). *Technological Trends in Epidemic Intelligence for Infectious Disease Surveillance: A Systematic Literature Review*. [In progress]
[11] Kasamatsu, A., Ota, M., Shimada, T., Fukusumi, M., Yamagishi, T., Samuel, A., Nakashita, M., Ukai, T., Kurosawa, K., Urakawa, M., Takahashi, K., Tsukada, K., Futami, A., Inoue, H., Omori, S., Kobayashi, M., Komiya, H., Shimada, T., Tabata, S., … Saito, T. (2021). Enhanced event-based surveillance for imported diseases during the Tokyo 2020 Olympic and Paralympic Games. *Western Pacific Surveillance and Response Journal*, *12*(4). https://doi.org/10.5365/wpsar.2021.12.4.903
[12] Lencucha, R., & Bandara, S. (2021). Trust, risk, and the challenge of information sharing during a health emergency. *Globalization and Health*, *17*(1). https://doi.org/10.1186/s12992-021-00673-9
[13] MacIntyre, C. R., Chen, X., Kunasekaran, M., Quigley, A., Lim, S., Stone, H., Paik, H. Y., Yao, L., Heslop, D., Wei, W., Sarmiento, I., & Gurdasani, D. (2023). Artificial intelligence in public health: the potential of epidemic early warning systems. In *Journal of International Medical Research* (Vol. 51, Issue 3). SAGE Publications Ltd. https://doi.org/10.1177/03000605231159335
[14] Mahussin, N., Mohd Jaapar, A., & Mustafa, L. A. (2021). Volatility of Technology and Healthcare Sectors Before and During Covid-19 Pandemic. *Malaysian Journal of Science Health & Technology*, *7*(2), 39–47. https://doi.org/10.33102/mjosht.v7i2.166
[15] Mboussou, F., Ndumbi, P., Ngom, R., Kamassali, Z., Ogundiran, O., Van Beek, J., Williams, G., Okot, C., Hamblion, E. L., & Impouma, B. (2019). Infectious disease outbreaks in the African region: overview of events reported to the World Health Organization in 2018. *Epidemiology and Infection*, *147*, e299. https://doi.org/10.1017/S0950268819001912
[16] Md Nadzri, M. N., Md Zamri, A. S. S., Singh, S., Sumarni, M. G., Lai, C. H., Tan, C. V., Aris, T., Mohd Ibrahim, H., Gill, B. S., Mohd Ghazali, N., Md Iderus, N. H., Lim, M. C., Ahmad, L. C. R. Q., Kamarudin, M. K., Ahmad, N. A. R., Tee, K. K., & Zulkifli, A. A. (2024). Description of the COVID-19 epidemiology in Malaysia. *Frontiers in Public Health*, *12*. https://doi.org/10.3389/fpubh.2024.1289622
[17] Mohamad Shariff, N. S., & Nur Hayani Izzati Abd Hamid. (2021). Consumers' Buying Behavior Towards Online Shopping During The Covid-19 Pandemic: An Empirical Study In Malaysia. *Malaysian Journal of Science Health & Technology*, *7*(2), 1–7. https://doi.org/10.33102/mjosht.v7i2.164
[18] Morgan, O., & Pebody, R. (2022). The WHO Hub for Pandemic and Epidemic Intelligence; supporting better preparedness for future health emergencies. In *Eurosurveillance* (Vol. 27, Issue 20). European Centre for Disease Prevention and Control (ECDC). https://doi.org/10.2807/1560-7917.ES.2022.27.20.2200385
[19] Ricks, P. M., Njie, G. J., Dawood, F. S., Blain, A. E., Winstead, A., Popoola, A., Jones, C., Li, C., Fuller, J., Anantharam, P., Olson, N., Walker, A. T., Biggerstaff, M., Marston, B. J., Arthur, R. R., Bennett, S. D., & Moolenaar, R. L. (2022). Lessons Learned from CDC's Global COVID-19 Early Warning and Response Surveillance System. *Emerging Infectious Diseases*, *28*(13), S8–S16. https://doi.org/10.3201/EID2813.212544
[20] Shausan, A., Nazarathy, Y., & Dyda, A. (2023). Emerging data inputs for infectious diseases surveillance and decision making. In *Frontiers in Digital Health* (Vol. 5). Frontiers Media SA. https://doi.org/10.3389/fdgth.2023.1131731
[21] Soliman Ashraf, Alyafei Fawziya, & Elalaily Rania. (2020). The death rate for covid-19 is positively associated with gross domestic products. *Acta Biomedica*, *91*(2), 224–225. https://doi.org/10.23750/abm.v91i2.9514
[22] Spagnolo, Luigi., AbdelMalik, Philip., Doherty, Brian., Fabbri, Marco., Linge, Jens., Marin Ferrer, Montserrat., Moussas, Charalambos., Osato, Francesco., Petrillo, Mauro., Poljanšek, Karmen., Schnitzler, Johannes., Vernaccini, Luca., & European Commission. Joint Research Centre. (2020). *Integration of the Epidemic Intelligence from Open Sources (EIOS) system and the INFORM suite : enhancing early warning with contextual data for informed decision making.*
[23] Syed Mohamad, S. N., Sharifah Fairuz Syed Mohamad, Shahrina Ismail, Fathima Begum Syed Mohideen, Fatin Ahza Rosli, & Nur Faraheen Abdul Rahman. (2021). Preparedness and Attributes of COVID-19 among Malaysian Public during the Movement Control Order. *Malaysian Journal of Science Health & Technology*, *7*(2), 23–29. https://doi.org/10.33102/mjosht.v7i2.170
[24] WHO. (2022). *Public health surveillance for COVID-19 Key points*. https://www.who.int/publications/i/item/WHO-2019-nCoV-SurveillanceGuidance-2022.2
[25] Worldometer. (2024, May 14). *Coronavirus (COVID-19) Mortality Rate*. https://www.worldometers.info/coronavirus/coronavirus-death-rate/#correct
[26] Yanagawa, M., Lorenzo, J. C., Fukusumi, M., Shimada, T., Kasamatsu, A., Ota, M., Nakashita, M., Kobayashi, M., Yamagishi, T., Samuel, A., Ukai, T., Kurosawa, K., Urakawa, M., Takahashi, K., Tsukada, K., Futami, A., Inoue, H., Omori, S., Komiya, H., … Olowokure, B. (2022). Use of Epidemic Intelligence from Open Sources for global event-based surveillance of infectious diseases for the Tokyo 2020 Olympic and Paralympic Games. *Western Pacific Surveillance and Response Journal*, *13*(3). https://doi.org/10.5365/wpsar.2022.13.3.959.