

Original Paper

# Consistency of Daily Number of Reported COVID-19 Cases in 191 Countries From 2020 to 2022: Comparative Analysis of 2 Major Data Sources

Han Liu<sup>1</sup>, MPH; Huiying Zong<sup>2</sup>, MPH; Yang Yang<sup>3</sup>, PhD; David C Schwebel<sup>4</sup>, PhD; Bin Xie<sup>1</sup>, BM; Peishan Ning<sup>1</sup>, PhD; Zhenzhen Rao<sup>1</sup>, PhD; Li Li<sup>1</sup>, PhD; Guoqing Hu<sup>1</sup>, PhD

<sup>1</sup>Department of Epidemiology and Health Statistics, Hunan Provincial Key Laboratory of Clinical Epidemiology, Xiangya School of Public Health, Central South University, Changsha, Hunan, China

<sup>2</sup>Department of Medical Records Management, Xiangya Hospital, Changsha, China

<sup>3</sup>Department of Statistics, Franklin College of Arts and Sciences, University of Georgia, Athens, GA, United States

<sup>4</sup>Department of Psychology, University of Alabama at Birmingham, Birmingham, AL, United States

## Corresponding Author:

Guoqing Hu, PhD

Department of Epidemiology and Health Statistics

Hunan Provincial Key Laboratory of Clinical Epidemiology, Xiangya School of Public Health, Central South University

172 Tongzipo Road, Yuelu District

Changsha, Hunan, 741000

China

Phone: 86 73189667218

Email: [huguoqing009@gmail.com](mailto:huguoqing009@gmail.com)

## Abstract

**Background:** The COVID-19 pandemic represents one of the most challenging public health emergencies in recent world history, causing about 7.07 million deaths globally by September 24, 2024. Accurate, timely, and consistent data are critical for early response to situations like the COVID-19 pandemic.

**Objective:** This study aimed to evaluate consistency of daily reported COVID-19 cases in 191 countries from the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE) and the World Health Organization (WHO) dashboards during 2020-2022.

**Methods:** We retrieved data concerning new daily COVID-19 cases in 191 countries covered by both data sources from January 22, 2020, to December 31, 2022. The ratios of numbers of daily reported cases from the 2 sources were calculated to measure data consistency. We performed simple linear regression to examine significant changes in the ratio of numbers of daily reported cases during the study period.

**Results:** Of 191 WHO member countries, only 60 displayed excellent data consistency in the number of daily reported COVID-19 cases between the WHO and JHU CSSE dashboards (mean ratio 0.9-1.1). Data consistency changed greatly across the 191 countries from 2020 to 2022 and differed across 4 types of countries, categorized by income. Data inconsistency between the 2 data sources generally decreased slightly over time, both for the 191 countries combined and within the 4 types of income-defined countries. The absolute relative difference between the 2 data sources increased in 84 countries, particularly for Malta ( $R^2=0.25$ ), Montenegro ( $R^2=0.30$ ), and the United States ( $R^2=0.29$ ), but it decreased significantly in 40 countries.

**Conclusions:** The inconsistency between the 2 data sources warrants further research. Construction of public health surveillance and data collection systems for public health emergencies like the COVID-19 pandemic should be strengthened in the future.

*JMIR Public Health Surveill* 2025;11:e65439; doi: [10.2196/65439](https://doi.org/10.2196/65439)

**Keywords:** COVID-19; pandemic; data consistency; World Health Organization; data quality

## Introduction

The COVID-19 pandemic represents one of the greatest global public health challenges of this century. According to World Health Organization (WHO) statistics, by September 24, 2024, COVID-19 led to over 7.04 million deaths globally [1]. Critical flaws in timely data collection and reporting have hindered and misled the prevention and control of the pandemic over the past 3 years [2]. For example, during the early phase of the pandemic, the cumulative number of COVID-19 cases in the United States released by the Johns Hopkins University dashboard quickly exceeded 1 million individuals by the end of April 2020, substantially higher than governmental statistics [3]. Such discrepancies created considerable chaos for US governmental responses to the COVID-19 pandemic [4]. To better respond to future pandemics and other global health crises, it is helpful to examine and summarize important lessons we learned.

One key aspect of public health crisis and pandemic preparedness is the need for high-quality data to guide decision-making, practice, and research about prevention and control. Furthermore, 2 open-access data sources were widely used in research and decision-making about the COVID-19 pandemic [5], the COVID-19 dashboards released by the WHO and the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE). The 2 data sources collected data in different ways and have been reported as having somewhat inconsistent data in several countries [6]. However, a comprehensive evaluation of the consistency between the 2 data sources has not been conducted across multiple countries and over the key time period of the pandemic, from 2020 to 2022.

Therefore, this study systematically examined the consistency of the number for daily reported COVID-19 cases in 191 countries from 2 data sources, WHO and JHU CSSE, from 2020 to 2022.

## Methods

### Database

Given their wide use and numerous citations, we selected the publicly accessible WHO COVID-19 dashboard and the JHU CSSE COVID-19 dashboard as the major sources to investigate data patterns concerning COVID-19 pandemic infection rates [6]. The WHO dashboard was created by the WHO. In January 2020, the WHO developed the case definition of COVID-19 infection and requested all member countries to report the number of daily reported cases through the International Health Regulations national focal points. It also collected COVID-19 daily cases and death counts from internet-based public sources (public dashboards and social media) as a supplement [7]. The JHU CSSE dashboard was created by Johns Hopkins University in early January 2020. It gathered data from publicly accessible sources referenced in a version-controlled README file on the repository [5].

For this study, we retrieved publicly accessible numbers of daily reported COVID-19 cases in the 191 United Nations member countries from both data sources between January 22, 2020 and December 31, 2022. In total, there were 6 days with missing values in the 2 data sources for all 191 countries and all study days together, comprising 0.00003% (6/205,325) of all study days. Considering the small number of missing values and the fact that they were unlikely to create significant impact on our results, we excluded those days from our analysis. Data before January 22, 2020, were also not included, as the outbreak had just initiated and the JHU CSSE COVID-19 dashboard was not yet established.

### Data Analysis

In total, 2 measures were used to assess consistency between the 2 data sources. One measure was the ratio of the daily number of reported cases from the JHU CSSE dashboard to that from the WHO dashboard, aligned by calendar time. As an example, in Afghanistan on March 31, 2022, the JHU CSSE dashboard reported 264 newly reported cases and the WHO dashboard reported 220. The ratio of JHU CSSE over WHO dashboards was calculated as “264/220=1.2” for that day in Afghanistan. To facilitate our description, we refer to the ratio of daily numbers of reported cases between the 2 data sources as “daily ratio” or, simply, “ratio.”

Based on the mean ratio of the 191 countries during the full study time period, we classified the daily ratios for each country into 5 categories of consistency: excellent ( $0.9 \leq \text{mean ratio} \leq 1.1$ ), high ( $0.8 \leq \text{mean ratio} < 0.9$  or  $1.1 < \text{mean ratio} \leq 1.2$ ), moderate ( $0.7 \leq \text{mean ratio} < 0.8$  or  $1.2 < \text{mean ratio} \leq 1.3$ ), low ( $0.4 \leq \text{mean ratio} < 0.7$  or  $1.3 < \text{mean ratio} \leq 1.6$ ), and poor (mean ratio  $< 0.4$  or  $1.6 < \text{mean ratio}$ ). Countries where the 2 dashboards had higher consistency yielded ratios closer to 1.0, of course, and countries where there were greater discrepancies between the 2 dashboards had ratios either far below or far above 1.0.

To address certain unusual cases, we empirically built the following rules to calculate the daily ratio in certain unusual cases: (1) when the numerator and denominator were both 0, we defined the ratio as 1, as this represented perfect agreement between the dashboards; (2) when only the numerator was 0 and the denominator was 10 or greater, we defined the ratio as 0.5, and when the denominator was less than 10, we used the fourth criterion listed below to define the value of the ratio; (3) when the denominator was 0 and numerator was 10 or greater, we defined the ratio as 2.0, and when the numerator was less than 10, we used the fourth criterion below to define the value of the ratio; and (4) when the numerator and denominator were both less than 10 and their differences were less than 3 (explaining 0.1% of all data), between 4 and 6 (explaining 0.3%), or  $\geq 7$  (explaining 0.5%), we defined the ratio as 0.9 or 1.1, 0.7 or 1.3, 0.5 or 1.5, respectively, thus avoiding extremely large or small ratios due to small numerators or denominators. We mapped the mean and coefficient of variation (CV) of the daily ratios between the 2 data sources for each of the 191 countries.

We graphed box plots and performed the Kruskal-Wallis rank sum test to display and compare differences in the mean and CV of the daily ratios across groups of countries defined by income levels. The 191 countries were classified into 4 income-based categories according to the World Bank Analytical Classifications of 2023 [8]: low-income countries, lower middle-income countries, upper middle-income countries, and higher-income countries. In addition, simple linear regression, with the absolute relative difference in daily number used as the dependent variable and time (date) as the independent variable, was performed to examine trends in the daily differences during the study period. A significant and positive regression coefficient indicates worsening data consistency over time, while a significant and negative regression coefficient indicates improved data consistency.

SPSS statistics 26 (IBM Corp), R version 4.3.0 (R Core Team), and Microsoft Excel spreadsheet 2021 were used for data analysis.  $P < .05$  was regarded as statistically significant.

### Ethical Considerations

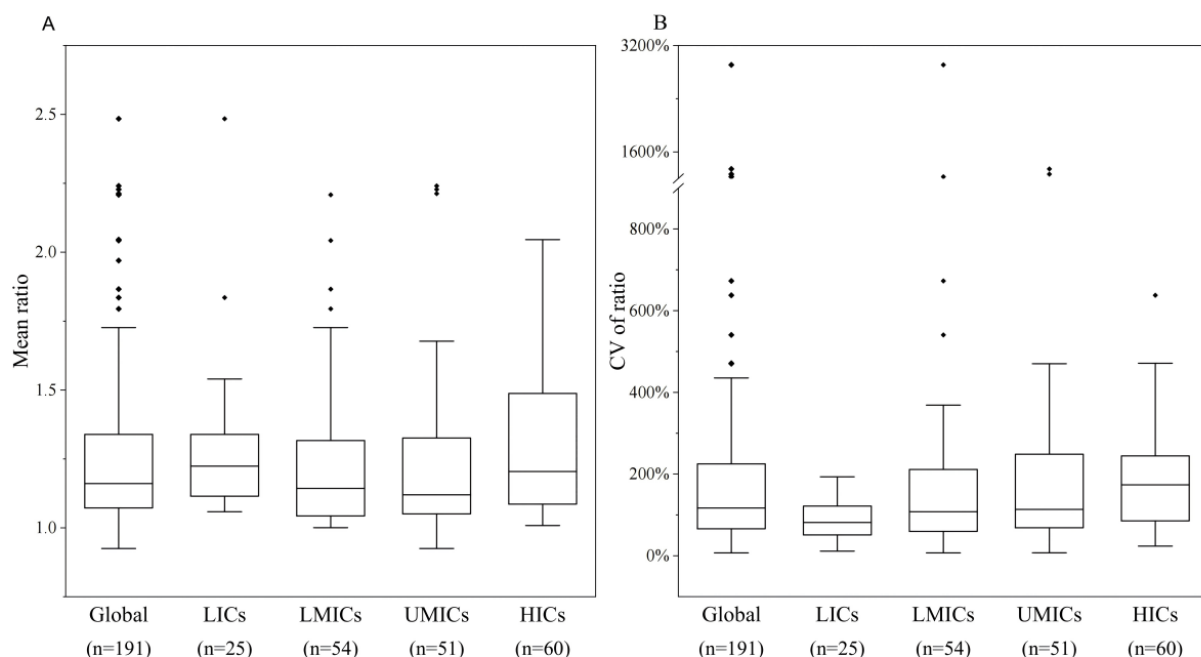
This study does not cause harm to humans and does not involve individual information or commercial interest

information. The data used in this study cannot be used to identify individual information, either directly or indirectly. Based on the given declarations, this study complies with the fundamental principles expressed in the Declaration of Helsinki and meets the ethical exemption requirements of the Ethical Review Measures for Life Sciences and Medical Research Involving Humans promulgated by China. Therefore, this research was exempted from ethical review.

## Results

For all observed dates combined, the median ratio between the daily number of reported COVID-19 cases in the 2 data sources for the 191 countries was 1.14 (IQR 1.07-1.34), with the lowest mean ratio in Turkey (0.92) and the highest mean ratio in Sudan (2.48) (Figure 1A). The ratios did not differ significantly across the 4 categories of countries organized by income ( $P = .12$ ). The median CV of ratio between the daily number of reported COVID-19 cases from the 2 data sources was 1.16 (IQR 0.66-2.25) for the 191 countries. The 4 types of countries had significantly different CVs of ratio ( $H = 10.08, P = .02$ ) (Figure 1B).

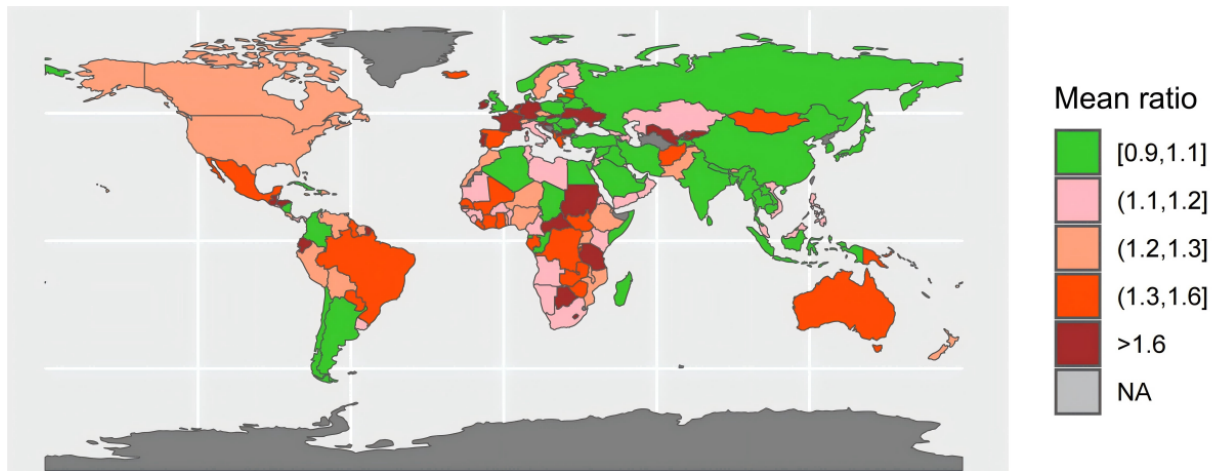
**Figure 1.** Ratio of number of daily new cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard and from the World Health Organization COVID-19 dashboard by country income category, from January 22, 2020, to December 31, 2022. CV: coefficient of variation; HIC: higher-income country; LIC: low-income country; LMIC: lower middle-income country; UMIC: upper middle-income country.



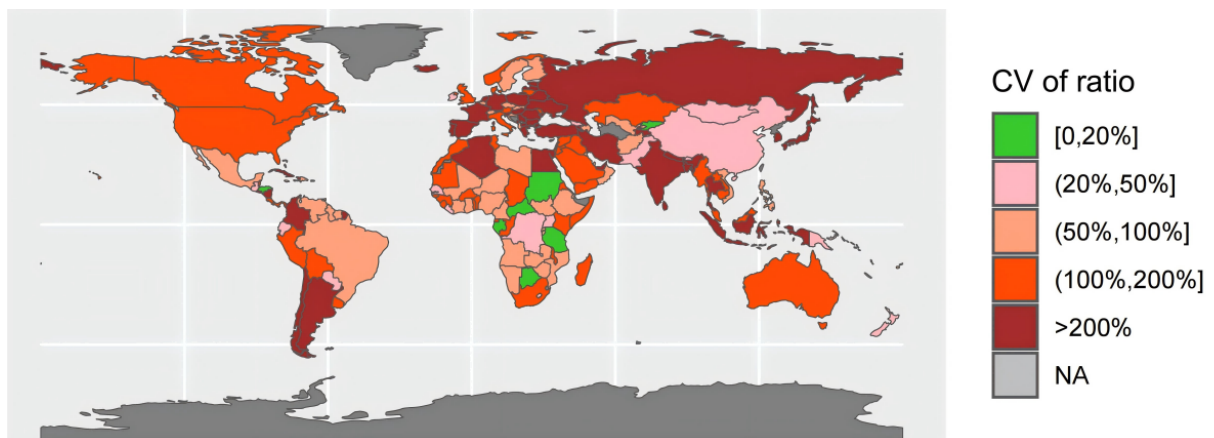
In Figure 1, the ratio was calculated as the number of daily new cases from the JHU CSSE COVID-19 dashboard divided by the number from the WHO COVID-19 dashboard, for each country. Figure 2 and Figure 3 display the variations in mean and CV of ratio between the daily number of reported COVID-19 cases from the 2 data sources across the countries. Of the 191 countries, 60 had excellent consistency (mean ratio: 0.9-1.1), while 37 had low consistency (mean ratio: 0.4-0.7 or 1.3-1.6) and 23 had poor consistency (mean

ratio:  $< 0.4$  or  $> 1.6$ ) (Figure 2). Most African and European countries had low or poor data consistency and most Asian countries had high or excellent data consistency. Strikingly, the CV of ratio between the number of daily reported COVID-19 cases from the 2 data sources during the study time period was higher than 100% in 110 countries (110/191, 57%), and 20% or less in only 8 countries (8/191, 4%) (Figure 3).

**Figure 2.** Mean ratio of number of daily new cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard and from the World Health Organization COVID-19 dashboard across 191 countries from January 22, 2020, to December 31, 2022. The ratio was calculated as the number of daily new cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard divided by the number from the World Health Organization COVID-19 dashboard, for each country. NA: not available.



**Figure 3.** CV of ratio of number of daily new cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard and from the World Health Organization COVID-19 dashboard across 191 countries from January 22, 2020, to December 31, 2022. The ratio was calculated as the number of daily new cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard divided by the number from the World Health Organization COVID-19 dashboard, for each country. CV: coefficient of variation; NA: not available.



For the 191 countries combined and within all 4 income-defined types of countries, the absolute relative difference between daily number of reported COVID-19 cases from the 2 data sources all increased slightly over the study time period, with  $R^2$  values of 0.16, 0.08, 0.06, 0.14, and 0.14, respectively (Table 1). Country-specific analyses showed that

the absolute relative difference increased in 84 countries but decreased in 40 countries significantly (Table 2). Notably, data inconsistency between the 2 data sources increased significantly over time in 3 countries (Malta:  $R^2=0.25$ ; Montenegro:  $R^2=0.30$ ; United States:  $R^2=0.29$ ).

**Table 1.** Linear trends in ratio of number of daily reported cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard and from the World Health Organization COVID-19 dashboard from January 22, 2020, to December 31, 2022, by country income category (dependent and independent variables of linear regression model were “the absolute value of (ratio-1)” and “time (date),” respectively).

Country income	Linear regression model	Coefficient of determination ( $R^2$ )	$P$ value
Global	$\hat{y} = -713.80 + 0.00029x$	0.16	<.001
LICs <sup>a</sup>	$\hat{y} = -1110.18 + 0.00045x$	0.08	<.001
LMICs <sup>b</sup>	$\hat{y} = -1025.15 + 0.00041x$	0.06	<.001
UMICs <sup>c</sup>	$\hat{y} = -781.22 + 0.00031x$	0.14	<.001
HICs <sup>d</sup>	$\hat{y} = -1453.78 + 0.00059x$	0.14	<.001

<sup>a</sup>LIC: low-income country.

<sup>b</sup>LMIC: lower middle-income country.



<sup>c</sup>UMIC: upper middle-income country.

<sup>d</sup>HIC: higher-income country.

**Table 2.** Linear trends in ratio of country-specific number of daily new cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard and from the World Health Organization COVID-19 dashboard from January 22, 2020, to December 31, 2022. Based on coefficient of determination ( $R^2$ ) of linear regression models with statistical significance, we classified linear trends of ratio into 3 grades: substantially ( $R^2 \geq 0.25$ ), moderately ( $0.10 < R^2 < 0.25$ ), and slightly ( $R^2 \leq 0.10$ ). In addition, we did not detect significant changes in data consistency for 67 countries).

Linear trend	Country
<b>Increased (n=84)</b>	
Substantially (n=3)	Malta ( $R^2=0.25$ ), Montenegro ( $R^2=0.30$ ), United States ( $R^2=0.29$ )
Moderately (n=4)	Australia ( $R^2=0.18$ ), Bulgaria ( $R^2=0.11$ ), Iran ( $R^2=0.13$ ), Slovenia ( $R^2=0.21$ )
Slightly (n=77)	Albania, Algeria, Antigua and Barbuda, Azerbaijan, Bahamas, Bangladesh, Barbados, Belgium, Bhutan, Bolivia, Bosnia and Herzegovina, Brazil, Brunei, Burundi, Cambodia, Canada, Chile, Comoros, Croatia, Cyprus, Czechia, Democratic Republic of Congo, Denmark, Dominica, Dominican Republic, Ecuador, Ethiopia, Fiji, Finland, Georgia, Grenada, Haiti, India, Iraq, Ireland, Italy, Kazakhstan, Kenya, Kiribati, Kuwait, Laos, Latvia, Liechtenstein, Lithuania, Maldives, Marshall Islands, Mauritius, Mexico, Micronesia, Monaco, Mongolia, Morocco, Namibia, Nauru, Norway, Palau, Panama, Papua New Guinea, Peru, Romania, Saint Kitts and Nevis, Samoa, San Marino, Saudi Arabia, Seychelles, Slovakia, Solomon Islands, South Africa, Sweden, Switzerland, Thailand, Timor, Tonga, Ukraine, United Kingdom, Vanuatu, Venezuela
<b>Decreased (n=40)</b>	
Moderately (n=2)	Poland ( $R^2=0.17$ ), Spain ( $R^2=0.11$ )
Slightly (n=38)	Andorra, Argentina, Bahrain, Belarus, Burkina Faso, Cameroon, Chad, Congo, Djibouti, Egypt, El Salvador, Estonia, Germany, Greece, Guinea, Hungary, Japan, Jordan, Madagascar, Malaysia, Mali, Mauritania, Moldova, Myanmar, Netherlands, North Macedonia, Portugal, Russia, Sierra Leone, Somalia, Togo, Tunisia, Turkey, Tuvalu, United Arab Emirates, Uruguay, Yemen, Zambia
No significant change (n=67)	Afghanistan, Angola, Armenia, Austria, Belize, Benin, Botswana, Cape Verde, Central African Republic, China, Colombia, Costa Rica, Cote d'Ivoire, Cuba, Equatorial Guinea, Eritrea, Eswatini, France, Gabon, Gambia, Ghana, Guatemala, Guinea-Bissau, Guyana, Honduras, Iceland, Indonesia, Israel, Jamaica, Kyrgyzstan, Lebanon, Lesotho, Liberia, Libya, Luxembourg, Malawi, Mozambique, Nepal, New Zealand, Nicaragua, Niger, Nigeria, Oman, Pakistan, Paraguay, Philippines, Qatar, Republic of Korea, Rwanda, Saint Lucia, Saint Vincent and the Grenadines, Sao Tome and Principe, Senegal, Serbia, Singapore, South Sudan, Sudan, Suriname, Sri Lanka, Syria, Tajikistan, Tanzania, Trinidad and Tobago, Uganda, Uzbekistan, Vietnam, Zimbabwe

## Discussion

### Main Findings

This study systematically examined the consistency of number of daily reported COVID-19 cases and number-based epidemic cycles in 191 countries from the 2 most cited data sources between 2020 and 2022. We generated 2 key findings. First, there was a difference in global number of daily reported COVID-19 cases between the WHO and JHU CSSE data, and data consistency differed between the 4 categories of income-defined countries and across the 191 countries. Second, data inconsistency between the 2 data sources generally decreased slightly over time, both for the 191 countries combined and within the 4 types of income-defined countries, but it changed differently over time across the 191 individual countries.

### Interpretation of Findings

Inconsistencies in the number of daily reported COVID-19 cases between the WHO and JHU CSSE data likely reflect differences in data collection and data release strategies across the 2 data sources. The WHO data came primarily from official daily counts reported by WHO member states, territories, and areas [7], while the JHU CSSE data were

derived from a combination of more than 400 sources for over 3500-point locations [5]. Compared with the WHO data, the JHU CSSE data were collected from a higher number of sources, potentially explaining why the number of daily reported COVID-19 cases from the JHU CSSE data was greater than that from the WHO data for the whole world, the 4 types of income-defined countries, and most individual countries during most dates [9]. The 2 data sources also adopted different data release strategies. The JHU CSSE COVID-19 dashboard used automated methods to extract data from each source every half hour and therefore updated data hourly [5]. In contrast, the WHO COVID-19 dashboard relied on official reports from all WHO member countries, territories, and areas, which were more prone to be delayed and less frequently updated than the JHU dashboard data [10].

Variations in data inconsistency between the WHO and JHU CSSE data might also be due to diverse data publishing policies and the evolution of data-releasing agencies (eg, delayed reporting, data entry errors, and data collection strategies), particularly at the local level across the 191 countries. As the pandemic eased, some countries like Sweden reduced the frequency of data release and even stopped publishing epidemic data, particularly at the local level [11]. These practices likely affected the original data

collected by the WHO and the JHU CSSE COVID-19 dashboards in different ways. In particular, sources indexed by the JHU CSSE dashboard were more extensive and were more likely to change over time than those covered by the WHO dashboard [12].

### **Implications and Limitations**

Our findings have 2 important implications. First, we do not have clear empirical basis to explain the drivers behind inconsistencies in the number of COVID-19 cases between the WHO and JHU CSSE dashboards or the changes in data inconsistency over time. A variety of internal and external factors likely could have contributed to the magnitude and pattern of the inconsistency, including but not limited to changes in government intervention policies or testing capacities.

Policy makers, researchers, and public health practitioners should recognize that the 2 data sources generate somewhat different and even occasionally conflicting results when they use the historical COVID-19 data collected by the WHO and JHU CSSE dashboards to summarize pandemic control experiences and lessons, conduct historical trend analyses, and evaluate intervention strategies [13-15]. Because strong evidence and detailed guidance are lacking concerning which of the 2 data sources is more valid for which specific time periods and countries, we recommend users consider both, recognizing each may have flaws. It also would be valuable to conduct further research assessing the quality of the data sources. For example, comparisons of the data sources with known cycles of the COVID-19 epidemic may yield additional information about the validity of each data source [16,17]. Another useful approach to bridge the gaps in high-level data from mainstream sources is to cross-validate them with data provided by local health authorities whenever and wherever available. This may be particularly useful for data on hospitalization and on fatalities, which are less sensitive to testing policies, health care-seeking behavior, and reporting bias [18].

Second, our results highlight the importance of gathering consistent and accurate data to fight against future pandemics or other public health emergencies [19]. Inconsistent data can lead to undesired consequences in decision-making, research, and prevention efforts [20,21]. Development of standardized data collection and release protocols, along with infrastructure- and competency-building to help all countries gather accurate health data, will prepare us not just for current health situations but also for future pandemics and broad public health emergencies [22].

Our research is primarily limited by the unavailability of a gold standard criterion. We assessed the consistency of 2 data sources and cannot comprehensively assess the reliability or validity of either, or identify which might be more accurate. No alternative or recognized valid measure of COVID-19 infections or fatalities exists worldwide. In addition, due to the unavailability of detailed information about COVID-19 data collection methods or data release policies across the 191 countries, we cannot reasonably interpret exactly why data consistency variations emerged across the 191 countries or why there were significant increases or decreases in data consistency over time for some countries. Future research may consider innovative strategies to identify key factors contributing to the observed data inconsistencies and significant changes in data consistency. Hospitalization data might be considered as a surrogate gold standard, for example, to approximate accuracy of other data sources [23], perhaps providing valuable clues. Advanced data analysis strategies, such as data harmonization or deep machine learning, might also help uncover reasons behind the results we report and be useful to address data accuracy in future public health emergencies. Finally, evidence-based data collection and release guidelines should be established by the WHO or other governing bodies, perhaps supplemented with financial and technical support to those countries with limited resources to follow international guidelines. With strong international collaboration, we should be able to improve transparency, consistency and overall quality of data emerging from public health crises in the future.

### **Conclusions**

We report a moderate difference in global average daily number of reported COVID-19 cases between 2020 and 2022 in the WHO and JHU CSSE datasets. The COVID-19 data inconsistency changed significantly across the countries studied and across time, both globally and in 124 individual countries. Data users should be cautious in interpreting the data, given inconsistency across the 2 most commonly cited COVID-19 data sources. Further research is recommended to develop guidelines to support valid interpretation of existing COVID-19 data, explore possible reasons for discrepancies, and understand factors such as governmental policies and data collection systems that may lead to data inconsistency. With those efforts, society can be best prepared to conduct valid health-related data collection in response to future pandemics and public health emergencies.

---

### **Acknowledgments**

The study was supported by Philosophy and Social Science Foundation of China (number 20&ZD120).

### **Data Availability**

Both databases are freely accessible for the public. Throughout the data utilization process, we strictly adhered to the data usage requirements of the source databases, ensuring no modifications were made to the metadata. The attribution of the data sources has been explicitly stated in the manuscript.

## Conflicts of Interest

None declared.

## References

1. WHO COVID-19 dashboard. World Health Organization. URL: <https://data.who.int/dashboards/covid19/data?n=c> [Accessed 2023-12-31]
2. Nageshwaran G, Harris RC, Guerche-Seblain CE. Review of the role of big data and digital technologies in controlling COVID-19 in Asia: public health interest vs. privacy. *Dig Health*. 2021;7:20552076211002953. [doi: [10.1177/20552076211002953](https://doi.org/10.1177/20552076211002953)] [Medline: [33815815](https://pubmed.ncbi.nlm.nih.gov/33815815/)]
3. Pollard CA, Morran MP, Nestor-Kalinoski AL. The COVID-19 pandemic: a global health crisis. *Physiol Genomics*. Nov 1, 2020;52(11):549-557. [doi: [10.1152/physiolgenomics.00089.2020](https://doi.org/10.1152/physiolgenomics.00089.2020)] [Medline: [32991251](https://pubmed.ncbi.nlm.nih.gov/32991251/)]
4. Lessons learned from COVID-19 pandemic: virtual regional consultation with informal expert group, 19 October 2021. World Health Organization. URL: <https://www.who.int/publications/i/item/sea-whe-6> [Accessed 2025-01-02]
5. Dong E, Ratcliff J, Goyea TD, et al. The Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard: data collection process, challenges faced, and lessons learned. *Lancet Infect Dis*. Dec 2022;22(12):e370-e376. [doi: [10.1016/S1473-3099\(22\)00434-0](https://doi.org/10.1016/S1473-3099(22)00434-0)] [Medline: [36057267](https://pubmed.ncbi.nlm.nih.gov/36057267/)]
6. Alamo T, Reina D, Mammarella M, Abella A. Covid-19: Open-data resources for monitoring, modeling, and forecasting the epidemic. *Electronics (Basel)*. 2020;9(5):827. [doi: [10.3390/electronics9050827](https://doi.org/10.3390/electronics9050827)]
7. Allan M, Lièvre M, Laurenson-Schafer H, et al. The World Health Organization COVID-19 surveillance database. *Int J Equity Health*. Nov 23, 2022;21(Suppl 3):167. [doi: [10.1186/s12939-022-01767-5](https://doi.org/10.1186/s12939-022-01767-5)] [Medline: [36419127](https://pubmed.ncbi.nlm.nih.gov/36419127/)]
8. World Bank country and lending groups. The World Bank. URL: <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519> [Accessed 2025-01-28]
9. Wu Z, Huang NE, Long SR, Peng CK. On the trend, detrending, and variability of nonlinear and nonstationary time series. *Proc Natl Acad Sci U S A*. Sep 18, 2007;104(38):14889-14894. [doi: [10.1073/pnas.0701020104](https://doi.org/10.1073/pnas.0701020104)] [Medline: [17846430](https://pubmed.ncbi.nlm.nih.gov/17846430/)]
10. Zuo X, Chen Y, Ohno-Machado L, Xu H. How do we share data in COVID-19 research? A systematic review of COVID-19 datasets in PubMed Central articles. *Brief Bioinform*. Mar 22, 2021;22(2):800-811. [doi: [10.1093/bib/bbaa331](https://doi.org/10.1093/bib/bbaa331)] [Medline: [33757278](https://pubmed.ncbi.nlm.nih.gov/33757278/)]
11. Hu T, Guan WW, Zhu X, et al. Building an open resources repository for COVID-19 research. *Data Inf Manag*. Sep 1, 2020;4(3):130-147. [doi: [10.2478/dim-2020-0012](https://doi.org/10.2478/dim-2020-0012)] [Medline: [35382104](https://pubmed.ncbi.nlm.nih.gov/35382104/)]
12. Lindström M. The COVID-19 pandemic and the timing of government response: a comparison of four Nordic countries March–June 2020. *Scand J Public Health*. Jul 2023;51(5):754-758. [doi: [10.1177/14034948231171201](https://doi.org/10.1177/14034948231171201)] [Medline: [37118947](https://pubmed.ncbi.nlm.nih.gov/37118947/)]
13. Badker R, Miller K, Pardee C, et al. Challenges in reported COVID-19 data: best practices and recommendations for future epidemics. *BMJ Glob Health*. May 2021;6(5):e005542. [doi: [10.1136/bmjgh-2021-005542](https://doi.org/10.1136/bmjgh-2021-005542)] [Medline: [33958393](https://pubmed.ncbi.nlm.nih.gov/33958393/)]
14. Badr HS, Du H, Marshall M, Dong E, Squire MM, Gardner LM. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. *Lancet Infect Dis*. Nov 2020;20(11):1247-1254. [doi: [10.1016/S1473-3099\(20\)30553-3](https://doi.org/10.1016/S1473-3099(20)30553-3)]
15. Guerrier C, McDonnell C, Magoc T, Fishe JN, Harle CA. Understanding health care administrators' data and information needs for decision making during the COVID-19 pandemic: a qualitative study at an academic health system. *MDM Policy Pract*. 2022;7(1):23814683221089844. [doi: [10.1177/23814683221089844](https://doi.org/10.1177/23814683221089844)] [Medline: [35368410](https://pubmed.ncbi.nlm.nih.gov/35368410/)]
16. Dowell SF, Ho MS. Seasonality of infectious diseases and severe acute respiratory syndrome—what we don't know can hurt us. *Lancet Infect Dis*. Nov 2004;4(11):704-708. [doi: [10.1016/S1473-3099\(04\)01177-6](https://doi.org/10.1016/S1473-3099(04)01177-6)]
17. Chemaitelly H, Tang P, Hasan MR, et al. Waning of BNT162b2 vaccine protection against SARS-CoV-2 infection in Qatar. *N Engl J Med*. Dec 9, 2021;385(24):e83. [doi: [10.1056/NEJMoa2114114](https://doi.org/10.1056/NEJMoa2114114)] [Medline: [34614327](https://pubmed.ncbi.nlm.nih.gov/34614327/)]
18. Huang AA, Huang SY. Hospitalized COVID-19 patients with diabetes have an increased risk for pneumonia, intensive care unit requirement, intubation, and death: a cross-sectional cohort study in Mexico in 2020. *Health Sci Rep*. Apr 2023;6(4):e1222. [doi: [10.1002/hsr2.1222](https://doi.org/10.1002/hsr2.1222)] [Medline: [37081996](https://pubmed.ncbi.nlm.nih.gov/37081996/)]
19. Johnson B. Five discoveries about COVID-19 made since the public health emergency ended. *Nat Med*. Dec 2023;29(12):2974-2976. [doi: [10.1038/d41591-023-00084-w](https://doi.org/10.1038/d41591-023-00084-w)]
20. Kung S, Doppen M, Black M, et al. Underestimation of COVID-19 mortality during the pandemic. *ERJ Open Res*. Jan 2021;7(1):00766-02020. [doi: [10.1183/23120541.00766-2020](https://doi.org/10.1183/23120541.00766-2020)]
21. TC J, Barbieri M, et al. Differences in COVID-19 mortality: Implications of imperfect and diverse data collection systems. *Popul Health Metr*. 2021;76:35-72. [doi: [10.3917/popu.2101.0035](https://doi.org/10.3917/popu.2101.0035)]

22. Suthar AB, Schubert S, Garon J, Couture A, Brown AM, Charania S. COVID-19 case definitions, diagnostic testing criteria, and surveillance across the pandemic's 25 highest burden countries. medRxiv. Preprint posted online on May 14, 2021. [doi: [10.1101/2021.05.11.21257047](https://doi.org/10.1101/2021.05.11.21257047)]
23. Huang AA, Huang SY. Diabetes is associated with increased risk of death in COVID-19 hospitalizations in Mexico 2020: a retrospective cohort study. Health Sci Rep. Jul 2023;6(7):e1416. [doi: [10.1002/hsr2.1416](https://doi.org/10.1002/hsr2.1416)] [Medline: [37415678](https://pubmed.ncbi.nlm.nih.gov/37415678/)]

## Abbreviations

**CV:** coefficient of variation

**JHU CSSE:** Johns Hopkins University Center for Systems Science and Engineering

**WHO:** World Health Organization

*Edited by Amaryllis Mavragani; peer-reviewed by Radhwan Ibrahim, Samuel Y Huang; submitted 15.08.2024; final revised version received 13.11.2024; accepted 24.11.2024; published 06.02.2025*

*Please cite as:*

*Liu H, Zong H, Yang Y, Schwebel DC, Xie B, Ning P, Rao Z, Li L, Hu G*

*Consistency of Daily Number of Reported COVID-19 Cases in 191 Countries From 2020 to 2022: Comparative Analysis of 2 Major Data Sources*

*JMIR Public Health Surveill 2025;11:e65439*

URL: <https://publichealth.jmir.org/2025/1/e65439>

doi: [10.2196/65439](https://doi.org/10.2196/65439)

© Han Liu, Huiying Zong, Yang Yang, David C Schwebel, Bin Xie, Peishan Ning, Zhenzhen Rao, Li Li, Guoqing Hu. Originally published in JMIR Public Health and Surveillance (<https://publichealth.jmir.org>), 06.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on <https://publichealth.jmir.org>, as well as this copyright and license information must be included.