BMJ Open

Understanding COVID-19 reporting behaviour to support political decision-making: a retrospective cross-sectional study of COVID-19 data reported to WHO

Auss Abbood ,[1] Alexander Ullrich,[2] Luisa A Denkel[1]

To cite: Abbood A, Ullrich A, Denkel LA. Understanding COVID-19 reporting behaviour to support political decision-making: a retrospective cross-sectional study of COVID-19 data reported to WHO. *BMJ Open* 2023;**13**:e061717. doi:10.1136/bmjopen-2022-061717

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2022-061717).

Received 21 February 2022
Accepted 01 December 2022

Check for updates

[1]Centre for International Health Protection, Robert Koch Institute, Berlin, Germany
[2]Infectious Disease Epidemiology Department, Robert Koch Institute, Berlin, Germany

Correspondence to
Auss Abbood; abbooda@rki.de

ABSTRACT

Objective Daily COVID-19 data reported by WHO may provide the basis for political ad hoc decisions including travel restrictions. Data reported by countries, however, are heterogeneous and metrics to evaluate its quality are scarce. In this work, we analysed COVID-19 case counts provided by WHO and developed tools to evaluate country-specific reporting behaviours.

Methods In this retrospective cross-sectional study, COVID-19 data reported daily to WHO from 3 January 2020 until 14 June 2021 were analysed. We proposed the concepts of binary reporting rate and relative reporting behaviour and performed descriptive analyses for all countries with these metrics. We developed a score to evaluate the consistency of incidence and binary reporting rates. Further, we performed spectral clustering of the binary reporting rate and relative reporting behaviour to identify salient patterns in these metrics.

Results Our final analysis included 222 countries and regions. Reporting scores varied between −0.17, indicating discrepancies between incidence and binary reporting rate, and 1.0 suggesting high consistency of these two metrics. Median reporting score for all countries was 0.71 (IQR 0.55–0.87). Descriptive analyses of the binary reporting rate and relative reporting behaviour showed constant reporting with a slight ‘weekend effect’ for most countries, while spectral clustering demonstrated that some countries had even more complex reporting patterns.

Conclusion The majority of countries reported COVID-19 cases when they did have cases to report. The identification of a slight ‘weekend effect’ suggests that COVID-19 case counts reported in the middle of the week may represent the best data basis for political ad hoc decisions. A few countries, however, showed unusual or highly irregular reporting that might require more careful interpretation. Our score system and cluster analyses might be applied by epidemiologists advising policy makers to consider country-specific reporting behaviours in political ad hoc decisions.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ This study used a new approach for analysing reporting behaviour of countries based on COVID-19 case count data submitted to WHO.
- ⇒ Our analyses, including the score system, are easy to replicate since no strong computational resources are required and only publicly available data and open-source software libraries were used.
- ⇒ Clustering and the score system developed are strong simplifications. False reporting or testing-related metrics (such as test positivity rate) are not considered by our self-developed score.
- ⇒ Further limitations such as different case detection, definitions, testing strategies, reporting practices or lag times by country as well as processing time by WHO or retrospective data corrections may also apply.

INTRODUCTION

The current COVID-19 pandemic demonstrated the urgency of a new international system for pandemic preparedness and response.[1] Further, it stresses national health systems, governments, populations, economies and international health organisations to their limits and beyond.

According to the Independent Panel for Pandemic Preparedness and Response, ‘national responses by countries have been most effective where decision-making authority was clear and if there was capacity to coordinate efforts across actors […].’[1] Further, formal advisory structures with the ability to provide timely scientific advice as well as the willingness to act on those were necessary for a successful national COVID-19 management.[1] It turned out that early implementation of these measures was crucial to ensure effectiveness.[2–6] Beyond that, the worldwide spread of COVID-19 showed that the response to infectious diseases needs to be carried out across borders. In March 2020, the European Union (EU) closed its external and Schengen borders to decelerate the rapid spread of COVID-19.[7] Later, travel restrictions became more targeted by most countries. The EU, the UK, the USA and Germany, for instance, implemented ad hoc travel restrictions or recommendations for travel risk areas that were limited to areas or countries with increased risk of COVID-19 transmission and/or the occurrence of new variants of concern (VOC).[8–11] These COVID-19 risk areas were updated regularly (eg, weekly) according to specific definitions that were adapted to the epidemic situation.[8–11] In Germany, political stakeholders decided approximately every week which countries to add, to erase or to further observe for such a list of risk areas between June 2020 and May 2022.[12] As a national public health institute, we were consulted by the German government to support these decisions.[9] By our scientific mandate, we had the ambition that our advice should be as evidence based and data driven as possible. Global COVID-19 data were provided by institutions and organisations like the European Centre for Disease Prevention and Control (ECDC), Our World in Data or WHO.[13–15] We decided to use WHO data for our approach as they were updated most regularly, generated from official data sources and were supposed to be acceptable by most countries. The WHO headquarters curates and publicly shares (almost) daily data on the COVID-19 pandemic from its member states.[16] The designation of COVID-19 travel risk areas required weekly ad hoc evaluations of epidemiological indicators including 7-day incidences (calculated from WHO case counts), the occurrence of VOC, qualitative parameters (eg, reports from public health institutes within the countries or personal communication) and others. Our work was motivated by the following three facts:

First, WHO data provided an important basis for political decisions, but analyses on characteristics of these data were scarce. Data quality of COVID-19 case counts from different sources has been criticised and discussed since the early phase of the pandemic.[17–19] Unfortunately, systematic analyses of WHO data and tools for ad hoc analysis of country-specific reporting behaviours were lacking. Further, we noticed that some countries did not report any cases on certain days of the week. Thus, we wanted to analyse and understand COVID-19 case count data submitted to WHO as good as possible to optimise ad hoc interpretations of WHO data for political decision makers. In addition to the existing literature, this work provides an assessment and descriptive classification of WHO data including analyses on the reporting behaviour/characteristics. We did a cross-sectional observational study of case count data provided by WHO from January 2020 until June 2021. Based on these data, we described, characterised and analysed reporting properties and patterns by calculating reporting scores and performing spectral clustering analyses.

Second, it is well known that early implementation of COVID-19 measures was crucial to ensure effectiveness.[2–6] Travel restrictions are heavy-handed measures for pandemic containment that turned out to be most challenging. They require real-time evaluation of the local epidemic situation, travel volumes and the epidemic situation in other countries.[20] Thus, timeliness of data is crucial for political ad hoc decisions like travel restrictions.[20] Scientific advisers need robust and timely shared data as well as prediction models to prevent epidemic policies from lagging behind fast-changing epidemic dynamics. On the regional or national level, real-time COVID-19 decision-making may be supported by statistical methods like nowcasting.[21–23] Such nowcasting approaches mitigate reporting delays but usually require information on infection and reporting date, or use other secondary data.[22] Unfortunately, the infection date, which is required for nowcasting, does not exist for WHO data. Additionally, inferential techniques may not be reliable enough for strong political measures as the closing of borders. We wanted to provide an easy tool that allows timely evaluations of COVID-19 reporting behaviour by country without applying prediction models, nowcasting or requiring manual reviews of time series of COVID-19 case counts reported to WHO for each country. We developed a reporting score to support interpretation of reporting behaviour for better informed ad hoc decision-making including implementation of travel restrictions.

Third, COVID-19 data were available for most, but not all countries. Successful pandemic management requires appropriate governance within countries as well as internationally coordinated actions. Such efforts are highly dependent on the willingness of countries to cooperate, share data and knowledge as well as centralising information.[24] However, healthcare activities including international cooperativeness depend on the design of country's health institutions. These may be affected by institutional influences including federalism, electoral competition, constitutional designs or political ideologies.[24] Thus, this work should highlight the importance of an appropriate pandemic governance by all countries including publicly and timely shared data. We analysed WHO reporting behaviour of countries to investigate whether this demand was fulfilled by most WHO member states or not.

Many countries implemented drastic public health measures including school closures, restrictions of mass gatherings, lockdowns, border measures, quarantine of travellers arriving from affected countries and many others to mitigate the spread of COVID-19.[25] For the next rise of COVID-19 case counts and/or future pandemics, it is important to learn which of the public health measures implemented proved most effective. An increasing number of systematic reviews and meta-analyses have been published to address this important question.[3 25 26]

This work, however, focused on a methodological approach to improve understanding of data sources, reporting behaviour and obstacles to consider for better informed political ad hoc decisions. We did an observational cross-sectional study of COVID-19 case counts reported by the WHO headquarters from January 2020 until June 2021. We used these data to characterise, describe and analyse COVID-19 reporting properties and patterns by developing reporting scores and performing spectral clustering analyses.

DATA AND METHODS

WHO data

This is a retrospective observational cross-sectional study. In this work, we analysed COVID-19 data reported by 236 countries, territories or areas to WHO from 3 January 2020 until 14 June 2021. Data on case and death counts were used as provided by WHO on 14 June 2021.[15] Data sets generated for this study and raw data used are available at our GitHub repository.[27] WHO invites countries to use case definitions as provided by WHO.[28] The first case definition for 'human infection with novel coronavirus (nCoV)' was published by WHO as interim guideline in January 2020.[29] Since then, several updates of WHO COVID-19 case definitions have been released to adapt to the current evidence available.[28 30–33]

Global data are compiled through WHO region-specific dashboards and/or aggregated count data reported to the WHO headquarters daily. More detailed information on data sources, definitions of new case and death counts can be found on the WHO COVID-19 dashboard[16] (data sources). Briefly, countries report cumulative numbers of COVID-19 cases and deaths to WHO on a daily basis. National public health institutes or national ministries of health usually conduct this reporting. In Europe, a system called TESSy is used to report COVID-19 data by national public health institutes to ECDC.[34] Subsequently, ECDC reports data to WHO. This process differs by country or region. Countries may report data stratified on the federal level. For the WHO headquarters, however, these data are available only aggregated to national level. WHO records the daily new infections and deaths of its member states and several metrics derived from these two numbers. One important metric is the 7-day incidence rate that is defined as the number of new cases or new deaths of the last 7 days per 100 000 inhabitants.

WHO reports the following changes in the data collection processes:

> From the 31 December 2019 to the 21 March 2020, WHO collected the numbers of confirmed COVID-19 cases and deaths through official communications under the International Health Regulations (IHR, 2005), complemented by monitoring the official ministries of health websites and social media accounts. Since 22 March 2020, global data are compiled through WHO region-specific dashboards (see links below), and/or aggregate count data reported to WHO headquarters daily.[16]

Although WHO data contain information on COVID-19-related deaths and cases, in this work, we focused on confirmed COVID-19 cases. We expected COVID-19 case counts to provide a more robust database for analyses of reporting behaviours due to the following reasons: COVID-19-related deaths occur less frequently (compared with COVID-19 case counts), they occur with a delay of days to weeks after infection (and consequently lag behind the current epidemic situation) and do face challenges in the attribution of cause of death.[35] Data analyses were stratified by time (day of the week) and location (global, regional and country specific). Reporting behaviour was assessed by investigating two indicators: binary reporting rate and relative reporting behaviour.

Binary reporting rate

First, binary reporting rate was defined as the relative reporting frequency per weekday regardless of the number of reported cases, that is, whether reporting occurred or not. The binary reporting rate was calculated over the whole time series if not indicated otherwise. As an example, if a country reported case numbers larger than zero for all Mondays in the period under review, it would get a binary reporting rate of 1 for Mondays. More explicitly, we divided the number of Mondays with reporting in the data set—in this case all Mondays—with the number of all Mondays in the data set. If on Tuesdays, the same country only reported every second week, the country would get a binary reporting rate of 0.5 for Tuesdays. For this example, we assumed an even number of Tuesdays in the time series. Then, reporting every second week is identical with reporting for half of all Tuesdays. Again, being explicit, we divided the number of Tuesdays—in this case half of all Tuesdays—by the number of all Tuesdays. Thus, we obtained a value of 0.5 in this example. This process was repeated for all weekdays.

Subsequently, we compared the binary reporting rate with the 7-day incidence rate per weekday and WHO region to identify a possible association between high case numbers and high reporting rates. Seven-day incidence rates aggregated per region were calculated as the mean incidence per weekday over all countries of this region. Analogously, the binary reporting rate was aggregated by calculating the binary reporting rate per weekday per country first. Thereupon, the mean binary reporting rate per weekday was calculated for all countries of this region.

Finally, we performed spectral clustering—a multivariate statistics method—that clusters similar binary reporting rates to reduce all country-specific binary reporting rates to a limited number of clusters.[36]

Reporting score

Based on observations in the data that in part showed diverging incidences and binary reporting rate (online supplemental appendix 1), we developed a score to quantify the discrepancy between reporting rate and the epidemiological situation. Seven-day incidence rates were scaled to a range between 0 and 1 to make them comparable to the binary reporting rate. We accomplished this by applying a min-max normalisation to the mean incidence per weekday and for each country by using the respective minimum and maximum incidence rates for scaling. Thus, the discrepancy between reporting rate and epidemic situation can be quantified as the difference between the mean binary reporting rate and the mean 7-day incidence rates within a country. In consequence, scores close to 1 indicate a high reporting rate and a comparably small incidence. This could be observed in countries that reported very frequently even if the number of new cases was small. We assumed that most countries with values close or equal to 1 have a successful COVID-19 response and a 'high probability of a high reporting diligence'. Values below 0 indicated insufficient reporting frequencies given relatively high incidences. These cases might indicate a strong reporting delay or other difficulties in reporting and represented countries with a 'low probability of a high reporting diligence'. A reporting score equal or close to 0 can be observed, that is, when the binary reporting score matches the scaled incidence of a country. Such scores may be interpreted as 'medium probability of a high reporting diligence' that needs closer examination only with medium priority. If incidences were generally high among countries, a higher binary reporting rate would be needed to achieve a score of 0 and *vice versa*. We call this measure the *reporting score*. Thus, the reporting score was defined as country-specific measure within the range of −1 and +1 that is based on the normalised means of 7-day incidence rates and binary reporting rates. Values close or equal to 1 are interpreted as the optimum while low reporting rates might require closer examinations with medium (reporting scores equal or close to 0) or high priority (negative reporting scores). However, very high reporting rates might also be an indicator of false reporting (please see limitations for more details).

The reporting score in countries with no/few COVID-19 cases will be close to 0 as the binary reporting by definition will be close to 0 in this scenario. When excluding false reporting of no cases, countries reporting no or small numbers of cases should also be able to receive a perfect reporting score as there are no problems with reporting. For such countries, we applied an imputation on the reporting rate. If a country reported less than seven cases per week, which is the minimum number of cases to theoretically achieve a perfect binary reporting score, it could never obtain a perfect reporting score although it might be reporting reliably. In this case, we imputed one single new case on days where actually no new case was reported. We did this n times per week and country $\max\left\{0,\ 7-\sum_{i=1}^{7} c_{w,i}\right\}$ and c is the number of cases, w is the week and i is the weekday. Assuming, for example, a country that reported only two cases on a Monday and no cases during the rest of the week would get five imputations for this week, for example, from Tuesday to Saturday. A country that exceeds seven cases in sum for a week gets no imputation for this week. This imputation was applied for 32 (mostly very small) countries. We tested the success of the imputation by correlating the population size (log) and reporting score with a Spearman rank test.

Table 1 Listing of all introduced metrics

Metric name	Definition (per country)	Imputed
Binary reporting rate	Relative frequency in which more than zero cases were reported stratified by weekday.	–
Relative reporting behaviour	Proportion of reported cases per weekday by the sum of cases per week.	–
Reporting score	Difference between binary reporting rate and 7-day incidence rate over 1 week.	Imputed for weeks with less than seven newly reported cases.

This table lists all introduced metrics by name together with their definition when applied to a single country and the information if the metric involves imputation and how this imputation was applied.

Relative reporting behaviour

For the relative reporting behaviour, the fraction of cases reported each weekday was divided by the sum of all cases reported in the same week. For example, if a country reported 70 new cases in week 1 and 10 new cases each day in this week, the relative statistic of this country in week 1 is 1/7 for each day. These analyses were also made on global, regional and country levels. While the binary reporting rate helps us spotting reporting gaps, relative reporting behaviour helps us detecting reporting lags. A typical reporting lag occurs after the weekend and could, for example, make Tuesday the day with the highest number of newly reported cases per week but only due to a logistical and not an epidemiological phenomenon which is made visible by the relative reporting behaviour.

All scores introduced in this section are summarised in table 1.

Simulations of reporting delays to illustrate adverse effects

To motivate our scores and show adverse effects of delayed reporting, we produced several scenarios with simulated reporting delays using the WHO COVID-19 data set of this work.[27] Delayed reporting decreases timeliness which is an important data quality property and particularly important for COVID-19.[37 38]

The impact of the delay in our scenarios is measured by the relative difference between the reported case counts from the actual, non-delayed case counts. The higher their difference, the worse the impact of the delay. We picked three 14-day-long time series from the data set with low, medium and high numbers of reported cases. Details of this method are described in the online supplemental methods appendix.

Patient and public involvement

Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

Table 2 Comparison of the binary reporting rate and mean 7-day incidence rate stratified by WHO regions from 3 January 2020 until 14 June 2021

WHO region	Mean binary reporting rate	Mean 7-day incidence rate (number of new cases in the last 7 days per 100 000 inhabitants)
Europe	0.80	89
Eastern Mediterranean	0.80	48
South-East Asia	0.75	27
Africa	0.55	13
Americas	0.52	50
Western Pacific	0.48	13

The mean binary reporting rate is the rate at which new COVID-19 cases were reported to WHO while the mean 7-day incidence rate refers to the 7-day incidence rates (number of cases over the last 7 days per 100 000 inhabitants) of all countries per WHO region.

RESULTS

WHO data from 3 January 2020 until 14 June 2021 include COVID-19 cases and deaths reported by 236 countries, territories or areas. For our analyses, we removed the following countries due to lack of reporting: American Samoa; Cook Islands; Kiribati; Democratic People's Republic of Korea; Federated States of Micronesia; Nauru; Niue; Palau; Pitcairn; Saint Helena, Ascension and Tristan da Cunha; Tokelau; Tonga; Turkmenistan; and Tuvalu. This selection resulted in 222 countries, territories or areas eligible for our analyses.

Global

The global binary reporting rate stratified by weekdays showed a higher reporting rate during the middle of the week and lower rates at the beginning and the end of the week (online supplemental appendix 2). This pattern is also visible when we stratify for region or country. You can find an analysis on the binary reporting rate per weekday and WHO region in online supplemental appendix 3.

Binary reporting behaviour

Table 2 shows a comparison of the binary reporting rate and the 7-day incidence rates (number of cases reported in the last 7 days per 100 000 inhabitants) per region.

The comparison of the mean incidence rates and binary reporting rates within one region showed that a low incidence rate does not always match with low reporting rates and vice versa (table 2). The WHO region Americas, for instance, had the second lowest reporting rate while reporting the second highest incidence rates at the same time.

Country level

According to the score system we developed to quantify the discrepancy between reporting rate and outbreak situation, only few countries exhibit a negative value (indicating insufficient reporting). Median reporting score for all countries included was 0.71 (IQR 0.55–0.87). The minimum reporting score was −0.17, while the maximum reporting score was 1.0 (figure 1).

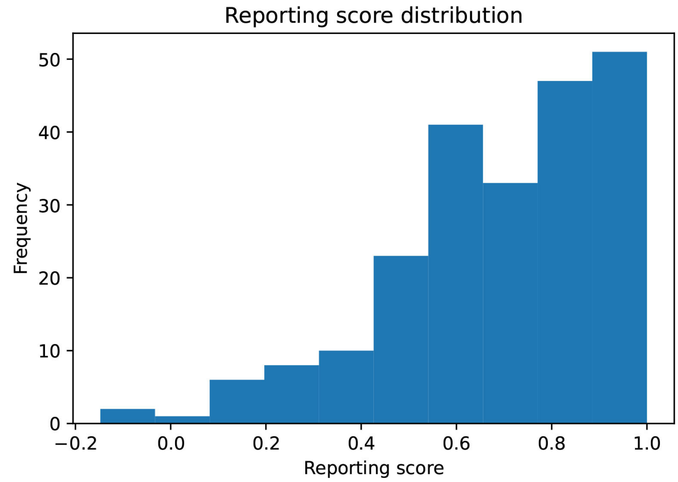


Figure 1 Worldwide distribution of the reporting score according to WHO data from 3 January 2020 until 14 June 2021. The reporting score is the difference of the binary reporting rate and the min-max scaled 7-day incidence rate. The 7-day incidence rate is the number of cases over the last 7 days per 100 000 inhabitants and the binary reporting rate quantifies the rate at which a country reported new cases of COVID-19. We applied imputation if a country reported less than 7 cases per week.

When excluding states with less than 500 000 inhabitants and an area less than 1000 km² which usually are islands with an advantage in mitigating the import of COVID-19 cases, we identified the highest scores for China (0.98), Tajikistan (0.97) and Egypt (0.97). The highest scores were interpreted as 'high probability of high reporting diligence' in most cases that required closer examination with only low priority. Even the lowest scores found in Montenegro (0.11), Czechia (0.13) and Slovenia (0.29) were in the positive range. These low positive scores might be interpreted as 'medium probability of high reporting diligence' that required closer examination with only medium priority.

A correlation of the population size and reporting score on the country level can be found in online supplemental appendix 4. The Spearman rank correlation coefficient for all countries is 0.28. To identify the outstanding patterns in a country's reporting behaviour, we applied spectral clustering to all 222 country/area-specific binary reporting rates and mapped them to six clusters (figure 2).

Figure 2 shows the following clusters: high constant reporting (cluster A; eg, Canada, Algeria and Nigeria), very low and sparse reporting (cluster B; eg, Greenland and Tanzania), lower constant reporting (cluster C; eg, Gibraltar and Cambodia), middle constant reporting with slight increases during the middle of the week (cluster D; eg, Iceland and Viet Nam), low reporting with slight increases during the middle of the week (cluster E; eg, Nicaragua and Congo) and few countries that rarely

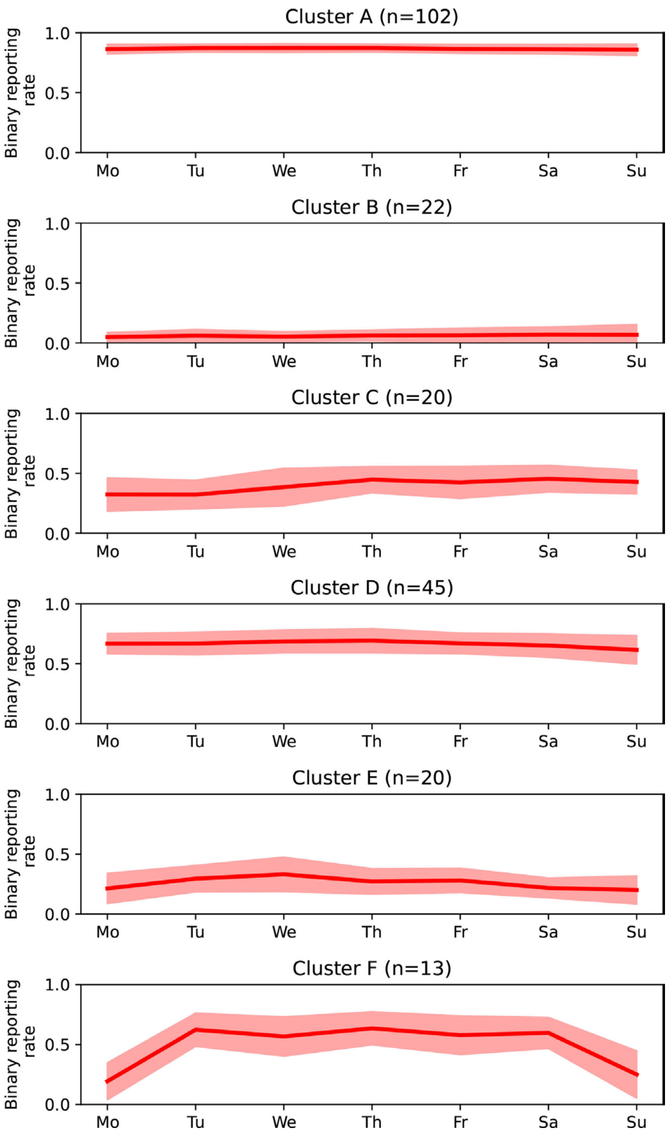


Figure 2 Clustering of binary reporting rates of countries according to WHO data from 3 January 2020 until 14 June 2021. The figure contains the result of clustering time series of binary reporting rates for each country using spectral clustering. The red lines are the mean reporting rate for each cluster with ±1 SD indicated by the red band. The title includes the number of countries that belong to each cluster.

report on Mondays or Sundays (cluster F; eg, Gabon, Norway and French Guiana).

In figure 2, clusters A and D showed a continuously high reporting rate while clusters B and C showed a constantly low reporting rate. Cluster F showed a mixed result. Therefore, we conducted another country-level analysis by counting the number of weekdays with a binary reporting rate below 50% (see online supplemental appendix 5). Using this count, we created a histogram visualising that most countries (n=137) have zero weekdays with a binary reporting rate below 50%. Interestingly, the second most frequent group of countries had a binary reporting rate of less than 50% for all weekdays (n=49). Consequently, 186 of 222 (84%) countries can be classified by only two groups as opposed to the clustering approach which used six clusters. The binary reporting rate tends to gravitate around the opposing extremes of the binary reporting behaviour.

This clustering result can be used in addition to the reporting score to identify countries with irregular reporting.

Our scenarios showed that currently observed case counts may be much higher or lower if reporting was delayed. The higher the deviation from the actual case counts, the larger the total number of cases and the longer the reporting delay. For the 25% quantile of reported cases within 14 days and a simulated reporting delay of 1 day, the deviation mainly ranged at ±100% with peaks close to 250%. If case counts were within the 75% quantile and there were 6 days of reporting delay, the difference from the actual case counts could range from around −300% up to 800% (online supplemental appendix 6). The median deviation decreased by quantile and the number of delayed days since we produced more zeros that caused a negative relative difference. The peaks in the 25% and 50% quantiles were higher because reporting in these quantiles was more regular than in the time series of the 75% quantile (online supplemental appendix 7).

Relative reporting behaviour

Global

The mean proportion of cases reported per week stratified by weekdays showed higher case numbers reported in the middle of the week and lower numbers reported at the beginning and the end of the week (see online supplemental appendix 8). This pattern matches the binary reporting rate that also found reporting gaps in the time from Sunday to Tuesday to occur most frequently. The analysis per WHO region can be found in online supplemental appendix 9.

We applied spectral clustering to identify patterns in relative reporting behaviour of 214 countries (figure 3). Eight countries could not be included due to non-finite values for their relative reporting behaviour score.

The majority of countries (n=135 countries) can be assigned to cluster A showing a straight line with a very modest increase in case counts in the middle of the week (Wednesday and Thursday).

The second largest cluster (cluster C, n=37) showed small peaks on Mondays and Wednesdays. This could be observed for several countries in Africa (eg, Botswana, Eritrea), America (eg, French Guiana, Dominica) but also Europe including Norway, Switzerland and Bulgaria. This clustering analysis helps identifying countries which are biased towards reporting a majority of their new cases on certain weekdays only. The smallest cluster (cluster E, n=5) showed a strong bias towards reporting cases on Sundays including countries like Nicaragua and Benin.

Sensitivity analyses

We conducted sensitivity analyses to evaluate the appropriateness of our imputation. Therefore, we plotted

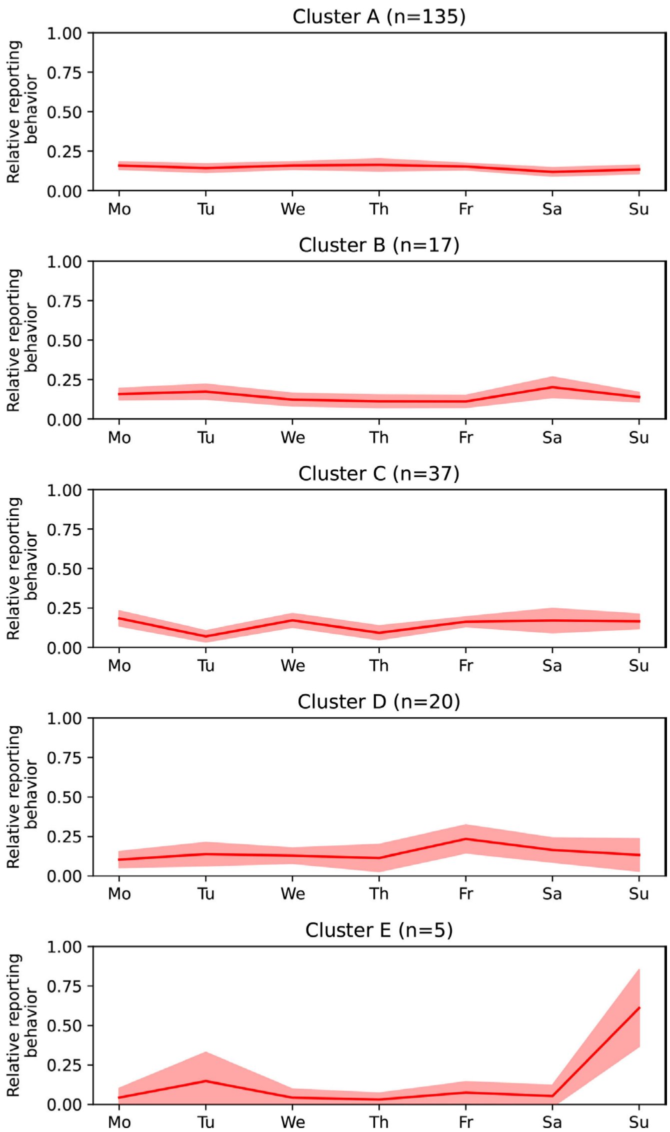


Figure 3 Clustering of the relative reporting behaviour of countries according to WHO data from 3 January 2020 until 14 June 2021. The figure contains the result of clustering time series of relative reporting behaviour for each country using spectral clustering. The red lines are the mean reporting rate of the cluster with ±1 SD indicated by the red band. The title includes the number of countries that belong to each cluster.

the distribution of the reporting score without imputation (online supplemental appendix 10), as well as data entries being removed if they would have been subject to imputation (online supplemental appendix 11). Our results confirmed the appropriateness of our imputation method.

Plotting the global binary reporting rate over time shows a low reporting rate that quickly increased around March 2020. The peak was reached in April 2020 suggesting a reliable reporting behaviour. A small decline over New Year's Day could be observed (online supplemental appendix 12). The great improvement of the reporting behaviour overlaps with the worldwide average 7-day incidence that also started to grow noticeable around March 2020 (online supplemental appendix 13). However, performing sensitivity analyses by excluding early data (January to March 2020) for calculation of country-specific reporting scores did not change our main findings (online supplemental appendix 14).

DISCUSSION

Our analyses showed that even though reporting behaviour of COVID-19 case counts by WHO member states was diverse, the majority of countries reported COVID-19 cases reliably and timely. Further, we developed a score system of reporting behaviour by country to support better informed political decisions. The median reporting score that can have a value between −1 and +1 was 0.71 for all countries. This median score suggests a relatively high consistency of incidences and binary reporting rates in many countries. The vast majority of countries showed constant reporting behaviour with a slight increase in case notifications during the middle of the week according to the patterns of the binary reporting rate. The relative reporting behaviour showed that a few countries report large numbers of cases on only a few days during the week.

Strengths and limitations

To our knowledge, this study used a new approach for analysing reporting behaviour of countries based on daily COVID-19 case counts submitted to WHO. This analysis is easy to replicate, no strong computational resources are required and only publicly available data and open-source software libraries were used. Our results were stable in all sensitivity analyses.

However, some limitations generally apply to analyses of public health data across countries with different resources and policies. These limitations were not identified by this work and continue to exist even with the results of the metrics introduced here. WHO reports these limitations in detail in the data source explanation.[15] First, these data underlie a variety of limitations including case detection, definitions, testing strategies, reporting practice and lag times (eg, time to case notification and time to reporting of deaths) that may differ between countries, territories and areas.[15] Second, processing time by WHO or retrospective corrections to data sets may have an impact on timeliness and accuracy of global case counts.[16] These factors, among others, may influence the counts leading to an underestimation or overestimation of true case and death counts. Thus, interpretation of data is difficult and needs to be performed carefully. According to WHO, COVID-19 counts primarily represent laboratory-confirmed cases and deaths as defined by WHO case definitions.[31] However, some differences may exist due to local adaptations. Difficulties and obstacles in interpreting and comparing COVID-19 case counts within and between countries have been discussed since the early phase of the pandemic.[17 18] Differences may occur due to regional and country-specific variations in testing capabilities, testing policies, case definitions and preparedness.[18 19] In consequence, it remains essential to consider additional indicators including testing rates, testing positivity, case fatality rates, hospitalisation, intensive care unit capacities and qualitative reports, for example, provided by embassies to gain a comprehensive picture of the COVID-19 situation in each individual country. Third, we developed a score to evaluate the quality of reporting behaviour by country. A score close to 1 indicated a 'high probability of high reporting diligence', while negative scores suggested a 'low probability of high reporting diligence' by the respective countries. However, this score system is a simplification and underlies limitations. Thus, interpretation of the country-specific scores must be done with caution. High scores should not be automatically interpreted as 'good reporting behavior', while low scores do not automatically represent 'poor reporting'. In this study, we interpret the frequency of reporting as a surrogate for accuracy and reliability of data. We used this assumption as a consequence of our calculated scenarios, existing literature and our experiences as epidemiological advisers for political ad hoc decisions during the COVID-19 pandemic.[22] Even if the pandemic situation was steady and case counts were low, most countries did report at least some single cases of COVID-19. However, for a few exceptions, this assumption might be incorrect. Less frequent reporting could also be an indicator of high data quality, as counts might be very thoroughly validated before reporting. Some countries or regions, in particular small island states, may in fact have no cases. Further, false reporting cannot be considered by our score system, but must be factored in for a comprehensive interpretation. Thus, high scores could also be achieved by consciously false reporting that we would not detect by our approach. However, we consider those as exceptional cases that need careful interpretation in the political context.

Patterns for binary reporting rate and relative reporting behaviour suggest that most countries reported more frequently and higher counts during the middle of the week. This might reflect structures of workweeks and working routines from Monday to Friday of general practitioners, laboratories, health authorities and public health institutes applied by most countries ('weekend effect'). In July 2021, WHO discontinued updating daily counts of COVID-19-confirmed cases and deaths on the dashboard at the weekend.[16] The term 'weekend effect' has been described as a suspected epidemiological effect causing substandard care during weekends due to higher workload and lower capacities of the healthcare system compared with workdays.[39] A 'weekend effect' has been described for other medical areas such as provision of emergency surgery or appropriate antibiotic prescription.[39–42] Higher reporting of COVID-19 case counts from Tuesday to Friday and day of week fluctuations have also been reported by regional studies and were attributed to temporal changes in laboratory testing and use of healthcare facilities as well as reporting delays.[43 44] Thus, for frequently repeated decision processes like ad hoc travel restrictions, it seems to be reasonable to use case counts reported in the middle of the week for ad hoc political decision-making to avoid underestimation of case counts. Some countries, however, showed extreme reporting behaviour such as reporting only once a week or reporting the majority of cases by the end of the week. Data from countries with extreme reporting behaviour must be interpreted with additional caution. Identifying them may be easier with our proposed metrics.

CONCLUSION

Global reporting behaviour of COVID-19 case counts by WHO member states was diverse, but the majority of countries reported COVID-19 cases when they did have cases to report. Furthermore, our clustering approach identified a 'weekend effect' suggesting COVID-19 case counts reported by WHO from the middle of the week being more reliable for advising political ad hoc decisions. Spectral clustering identified a few countries with unusual or irregular reporting that should be interpreted especially carefully. In consequence, elaborative manual review of WHO data time series and additional information could be restricted to countries with a 'low probability of high reporting diligence' and/or affiliation to certain clusters. However, our scores and cluster analyses should be applied keeping in mind its limitations. They do not replace thorough analyses of quantitative and qualitative indicators of the COVID-19 situation in each country for an informed decision-making.

Implication for epidemiologists advising policy makers

We developed a score system of WHO reporting behaviour that might be a helpful tool for infection control experts and epidemiologists advising policy makers. It may help them to consider country-specific reporting behaviours in political ad hoc decisions based on WHO data including designation of travel risk areas. The slight 'weekend effect' suggests that epidemiologists should prefer using COVID-19 case counts reported by WHO from the middle of the week, if possible, for advising political ad hoc decisions.

For the current and future pandemics, we need a robust system of epidemic intelligence to timely collect, share and analyse data at the regional, national and international levels for better informed political decisions. In September 2021, the first WHO hub for pandemic and epidemic intelligence was inaugurated in Berlin, Germany to achieve this ambitious goal.[45] The WHO hub aims to create a collaborative, interdisciplinary environment and may become the foundation for 'better data, better analytics and better decisions' for its 193 member states.[46] Successful pandemic management requires appropriate governance, but especially the willingness of countries to cooperate, share data and knowledge as well as centralising information. Finding ways to overcome these barriers will be a big challenge of our future.

Twitter Auss Abbood @Auss_Abbood

Contributors AA conducted all analyses, produced all figures and tables, was responsible for the conceptualisation of this work and wrote the initial draft of the manuscript. AU provided methodological supervision. LAD provided scientific supervision and was responsible for the conceptualisation of this work. All authors have read and revised the manuscript. AA acts as a guarantor for the overall content of this manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study does not involve human participants nor animal subjects. This research uses only publicly available records without contact with the individual/s. Thus, no ethical approval was required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Extra data can be accessed via the Dryad data repository at http://datadryad.org/ with the doi: 10.5061/dryad.9s4mw6mmb and from GitHub at https://github.com/aauss/reporting_behavior.

ORCID iD

Auss Abbood http://orcid.org/0000-0002-4428-168X

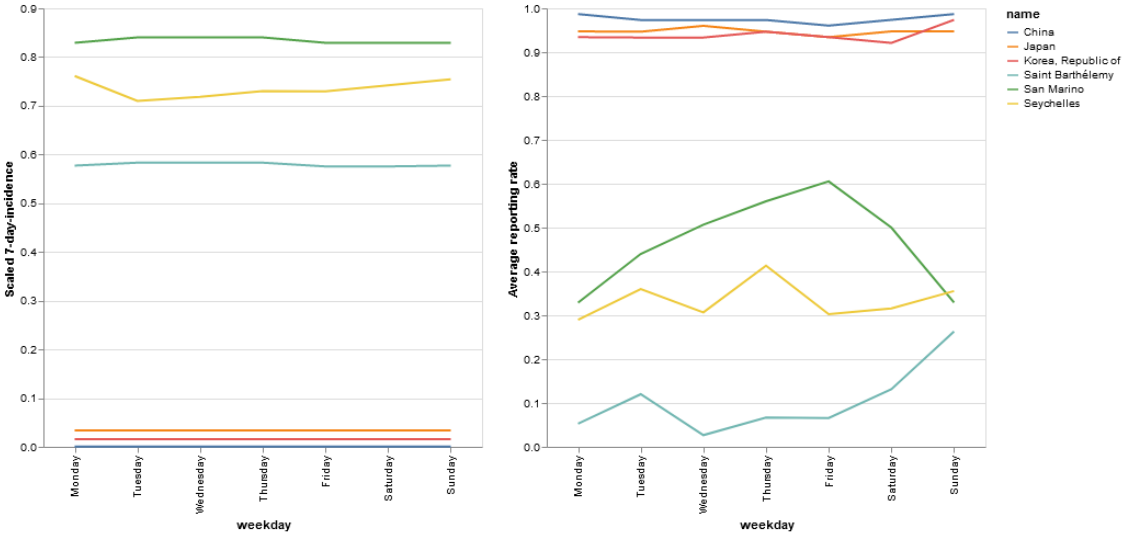
REFERENCES

1 The independent panel for pandemic preparedness and response. COVID-19: make it the last pandemic, 2021. Available: https://theindependentpanel.org/wp-content/uploads/2021/05/COVID-19-Make-it-the-Last-Pandemic_final.pdf
2 Askitas N, Tatsiramos K, Verheyden B. Estimating worldwide effects of non-pharmaceutical interventions on COVID-19 incidence and population mobility patterns using a multiple-event study. *Sci Rep* 2021;11:1972.
3 Iezadi S, Gholipour K, Azami-Aghdash S, *et al*. Effectiveness of non-pharmaceutical public health interventions against COVID-19: a systematic review and meta-analysis. *PLoS One* 2021;16:e0260371.
4 Meo SA, Abukhalaf AA, Alomar AA, *et al*. Impact of lockdown on COVID-19 prevalence and mortality during 2020 pandemic: observational analysis of 27 countries. *Eur J Med Res* 2020;25:56.
5 Thu TPB, Ngoc PNH, Hai NM, *et al*. Effect of the social distancing measures on the spread of COVID-19 in 10 highly infected countries. *Sci Total Environ* 2020;742:140430.
6 Zeng K, Bernardo SN, Havins WE. The use of digital tools to mitigate the COVID-19 pandemic: comparative retrospective study of six countries. *JMIR Public Health Surveill* 2020;6:e24598.
7 Rosik P, Komornicki T, Duma P, *et al*. The effect of border closure on road potential accessibility in the regions of the EU-27. The case of the COVID-19 pandemic. *Transp Policy* 2022;126:188–98.
8 European Centre for Disease Prevention and Control (ECDC). Maps in support of the Council recommendation on a coordinated approach to travel measures in the EU 2020, 2022. Available: https://www.ecdc.europa.eu/en/covid-19/situation-updates/weekly-maps-coordinated-restriction-free-movement [Accessed 18 Sep 2022].
9 Robert Koch Institute (RKI). Information on the designation of international risk area 2022, 2022. Available: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Transport/Archiv_Risikogebiete/Risikogebiete_2022-02-25_en.pdf?__blob=publicationFile [Accessed 16 May 2022].
10 United Kingdom Department for Transport UKDoHaSC. Guidance: travel to England from another country during coronavirus (COVID-19) 2021, 2022. Available: https://www.gov.uk/guidance/travel-to-england-from-another-country-during-coronavirus-covid-19#red-list-countries-and-territories [Accessed 18 Sep 2022].
11 United States (US) Centers for Disease Control and Prevention (CDC). How CDC determines the level for COVID-19 travel health notices 2022, 2022. Available: https://wwwnc.cdc.gov/travel/page/covid-19-travel-health-notice-levels [Accessed 18 Sep 2022].
12 Robert Koch Institute (RKI). Informationen Zur Ausweisung internationaler Risikogebiete durch das Auswärtige amt, BMG und BMI 2022, 2022. Available: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Risikogebiete_neu.html [Accessed 8 Nov 2022].
13 European Centre for Disease Prevention and Control (ECDC). COVID-19 situation update worldwide 2022. Available: https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases [Accessed 16 May 2022].
14 Our World in Data. Coronavirus pandemic (COVID-19), 2022. Available: https://ourworldindata.org/coronavirus [Accessed 16 May 2022].
15 World Health Organization. WHO coronavirus (COVID-19) dashboard, 2021. Available: https://covid19.who.int/info2021
16 World Health Organization. WHO coronavirus (COVID-19) dashboard - data sources, 2022. Available: https://covid19.who.int/data [Accessed 13 May 2022].
17 Jagodnik KM, Ray F, Giorgi FM. Correcting under-reported COVID-19 case numbers: estimating the true scale of the pandemic. *medRxiv*2020:20036178.
18 Odone A, Delmonte D, Scognamiglio T, *et al*. COVID-19 deaths in Lombardy, Italy: data in context. *Lancet Public Health* 2020;5:e310.
19 Suthar AB, Schubert S, Garon J, *et al*. Coronavirus disease case definitions, diagnostic testing criteria, and surveillance in 25 countries with highest reported case counts. *Emerg Infect Dis* 2022;28:148.
20 Russell TW, Wu JT, Clifford S, *et al*. Effect of internationally imported cases on internal spread of COVID-19: a mathematical modelling study. *Lancet Public Health* 2021;6:e12–20.
21 Günther F, Bender A, Katz K, *et al*. Nowcasting the COVID-19 pandemic in Bavaria. *Biom J* 2021;63:490–502.
22 Harris JE. Timely epidemic monitoring in the presence of reporting delays: Anticipating the COVID-19 surge in New York City, September 2020. *BMC Public Health* 2022;22:871.
23 Russell Timothy WHJ, Sam A, Christopher J. Using a delay-adjusted case fatality ratio to estimate under-reporting (Preprint). *Fondazione CERM* 2020.
24 Costa-Font J, Turati G, Batinti A. *The political economy of health and healthcare: the rise of the patient citizen*. Cambridge University Press, 2020.
25 Ayouni I, Maatoug J, Dhouib W, *et al*. Effective public health measures to mitigate the spread of COVID-19: a systematic review. *BMC Public Health* 2021;21:1015.
26 Nussbaumer-Streit B, Mayr V, Dobrescu AI, *et al*. Quarantine alone or in combination with other public health measures to control COVID-19: a rapid review. *Cochrane Database Syst Rev* 2020;9:Cd013574.
27 Abbood A, Ullrich A, Denkel LA. Understanding COVID-19 reporting behavior to support political decision-making: a retrospective cross-sectional study of COVID-19 data reported to the World Health Organization. Github, June 16, 2021, 2021. Available: https://github.com/aauss/reporting_behavior
28 World Health Organization. *WHO COVID-19: case definitions: updated in Public health surveillance for COVID-19, 22 July 2022*. Geneva: World Health Organization, 2022. https://apps.who.int/iris/handle/10665/360579 accessed WHO/2019-nCoV/Surveillance_Case_Definition/2022.1
29 World Health Organization. *Surveillance case definitions for human infection with novel coronavirus (nCoV), interim guidance, 15 January 2020*. Geneva: World Health Organization, 2020. https://apps.who.int/iris/handle/10665/332412 accessed WHO/2019-nCoV/Surveillance/2020.2
30 World Health Organization. *WHO COVID-19 case definition*. Geneva: World Health Organization, 2020. https://apps.who.int/iris/handle/10665/333912 accessed WHO/2019-nCoV/Surveillance_Case_Definition/2020.1
31 World Health Organization. *WHO COVID-19: case definitions: updated in public health surveillance for COVID-19*. Geneva: World Health Organization, 2020. https://apps.who.int/iris/handle/

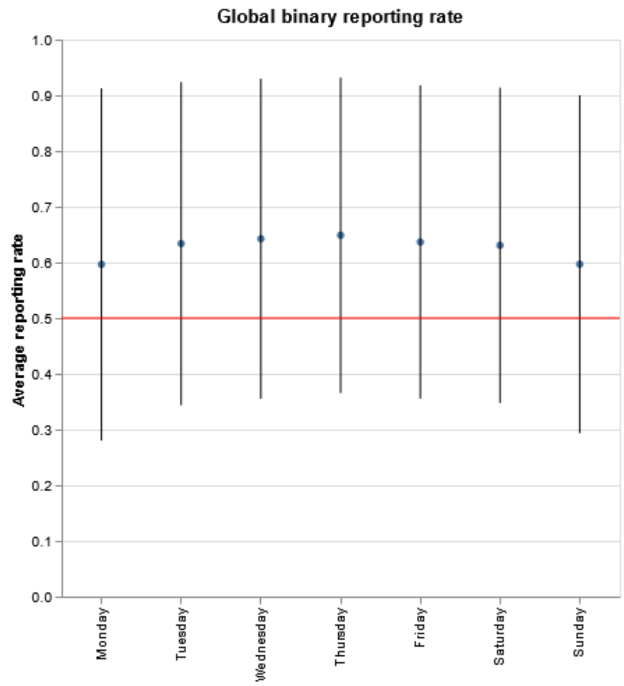
10665/337834 accessed WHO/2019-nCoV/Surveillance_Case_ Definition/2020.2
32 World Health Organization. Public health surveillance for COVID-19: interim guidance, 2020. Available: https://www.who.int/publications/i/item/who-2019-nCoV-surveillanceguidance-2020.8
33 World Health Organization. WHO COVID-19 case definition, 2020. Available: https://www.who.int/publications/i/item/WHO-2019-nCoV-Surveillance_Case_Definition-2020.2
34 European Centre for Disease Prevention and Control (ECDC). The European surveillance system (TESSy) 2022, 2022. Available: https://www.ecdc.europa.eu/en/publications-data/european-surveillance-system-tessy [Accessed 18 Sep 2022].
35 Battegay M, Kuehl R, Tschudin-Sutter S, *et al*. 2019-novel coronavirus (2019-nCoV): estimating the case fatality rate – a word of caution. *Swiss Med Wkly* 2020;5.
36 Shi J, Malik J. Normalized cuts and image segmentation. IEEE transactions on pattern analysis and machine intelligence 2000;22:888–905.
37 Lee YW, Wang RY, Wang RY. Data quality assessment. *Commun ACM* 2002;45:211–8.
38 Tucker Catherine WYC. The role of delayed data in the COVID-19 pandemic (Preprint). *SSM*2021.
39 Tebala GD, Milani MS, Cirocchi R, *et al*. The weekend effect on the provision of emergency surgery before and during the COVID-19 pandemic: case-control analysis of a retrospective multicentre database. *World J Emerg Surg* 2022;17:22.
40 Bishara J, Hershkovitz D, Paul M, *et al*. Appropriateness of antibiotic therapy on weekends versus weekdays. *J Antimicrob Chemother* 2007;60:625–8.
41 Hatchimonji JS, Kaufman EJ, Sharoky CE, *et al*. A 'weekend effect' in operative emergency general surgery. *Am J Surg* 2020;220:237–9.
42 Metcalfe D, Castillo-Angeles M, Rios-Diaz AJ, *et al*. Is there a "weekend effect" in emergency general surgery? *J Surg Res* 2018;222:219–24.
43 Greene SK, McGough SF, Culp GM, *et al*. Nowcasting for real-time COVID-19 tracking in New York City: an evaluation using reportable disease data from early in the pandemic. *JMIR Public Health Surveill* 2021;7:e25538.
44 Simpson RB, Lauren BN, Schipper KH, *et al*. Critical periods, critical time points and Day-of-the-Week effects in COVID-19 surveillance data: an example in Middlesex County, Massachusetts, USA. *Int J Environ Res Public Health* 2022;19:19031321. doi:10.3390/ijerph19031321
45 World Health Organization. WHO hub for pandemic and epidemic intelligence. Better data. Better analytics. Better decisions, 2022. Available: https://pandemichub.who.int/ [Accessed 18 Sep 2022].
46 World Health Organization. WHO hub for pandemic and epidemic intelligence. Better data. Better analytics. Better decisions 2021.

Appendix

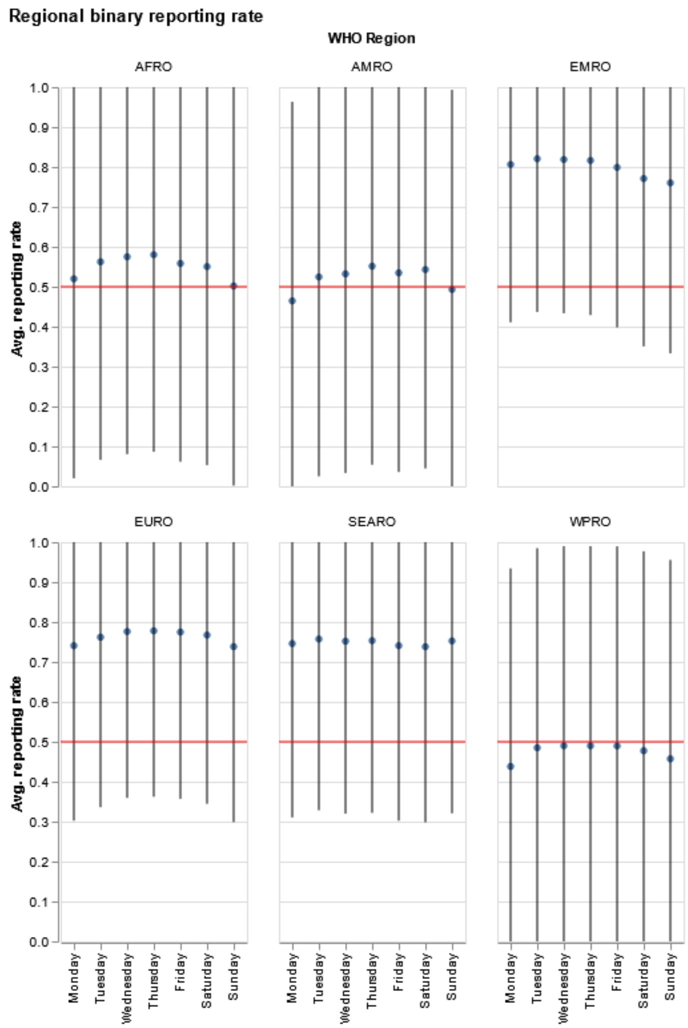
COVID-19 data reported daily to WHO from 3rd January 2020 until 14th June 2021 were analyzed in all figures if not indicated otherwise.



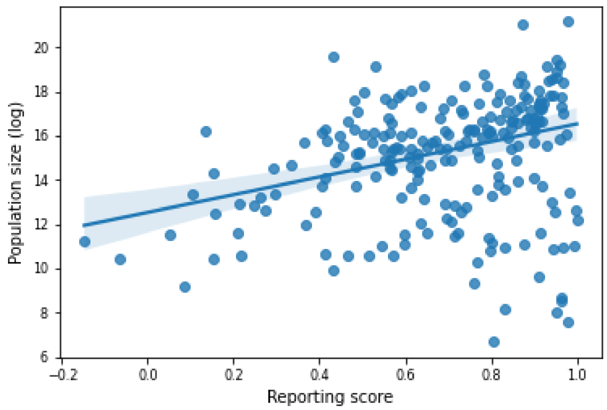
Appendix 1: Comparison of the scaled 7-day-incidence and the binary reporting rate for six countries. Countries were selected based on their minimal and maximal discrepancy between incidence and reporting rate.



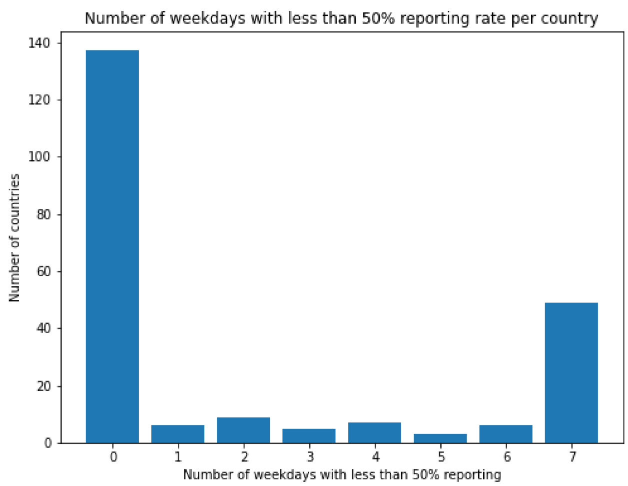
Appendix 2: Global reporting rate per weekday. This plot shows the binary reporting rate, i.e., the rate with which more than zero COVID-19 cases were reported to who by weekday averaged over all countries. The error bar around each dot indicates one standard deviation.



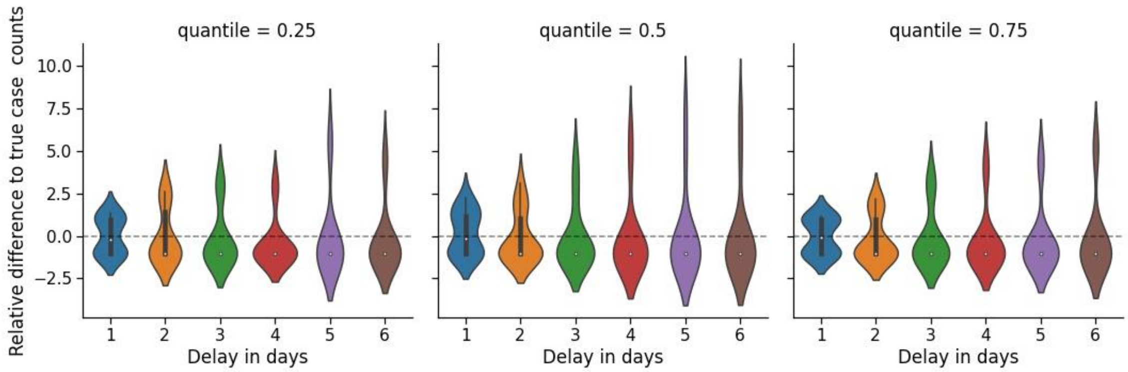
Appendix 3: Binary reporting rate by weekday and WHO region. This plot shows the binary reporting rate, i.e., the rate with which more than zero COVID-19 cases were reported to WHO by weekday and WHO region. The error bar around each dot indicate one standard deviation and the red horizontal lines indicate a 50% reporting rate. The abbreviations stand for the African (AFRO), American (AMRO), Eastern Mediterranean (EMRO), European (EURO), South-East Asian (SEARO), and West Pacific (WPRO) WHO regional offices.



Appendix 4: Correlation of population size and reporting score. This scatter plot represents all countries by their reporting score and (log) population size. The blue line is a linear fit of the data and the light blue band is the 95% confidence interval for the fit. The Spearman rank correlation coefficient is 0.28.



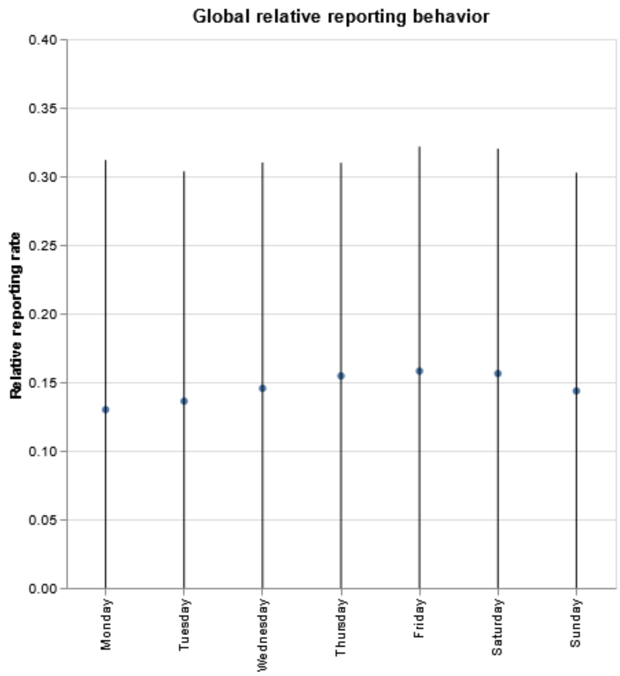
Appendix 5: Number of weekdays with less than 50% reporting rate per country. For each country, we counted the number of weekdays with binary reporting rates below 50%, i.e., the weekdays in which less than 50% of the weeks case counts above zero were reported.



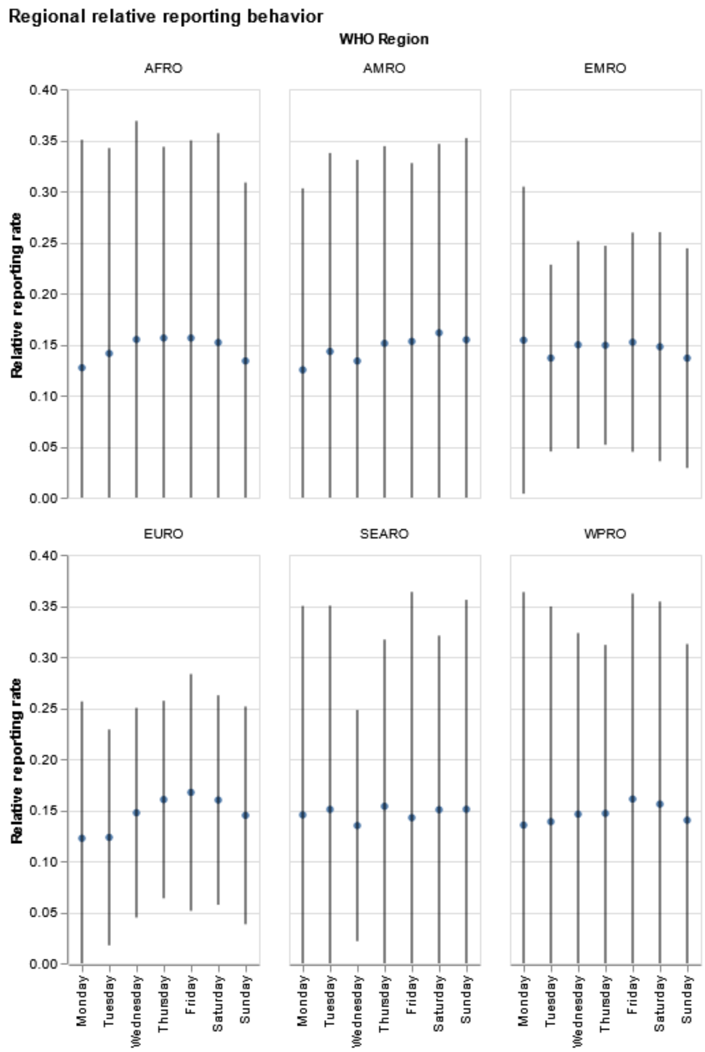
Appendix 6: Measurement of the inaccuracies a reporting delay causes using scenarios. For all countries, reported cases were split into 14-day long time series. Time series were summarized by their sum and the sums that were closest to the 25%, 50%, and 75% percentile were chosen for this analysis. Reporting delays were simulated by subtracting all cases in the delay period and add them to the next regular reporting day, e.g., with a two-day reporting day, a timeseries of [10,14,22,15,7] turns into [0,0,24,0,0,37]. The relative difference between the reported case counts and delayed case counts per day was measured and plotted for the different quantiles and different reporting delays. No difference, i.e., a relative difference of zero is marked with a dashed, black, horizontal line.

Appendix 7: Time series used for scenarios with artificial delays.

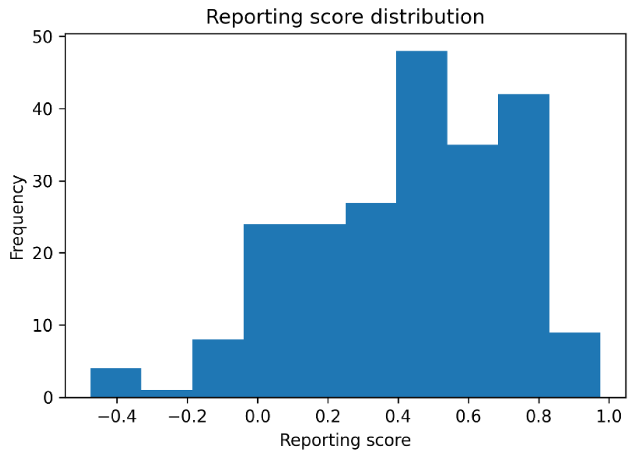
Quantiles	Reported case counts from day 1 to 14
25% percentile	89, 67, 65, 78, 89, 66, 90, 89, 81, 131, 134, 102, 179, 184
50% percentile	438, 268, 392, 336, 348, 397, 476, 632, 619, 344, 861, 390, 549, 460
75% percentile	1592, 1717, 1543, 1366, 1739, 1903, 1932, 1863, 1968, 1684, 1529, 1960, 2014, 2088



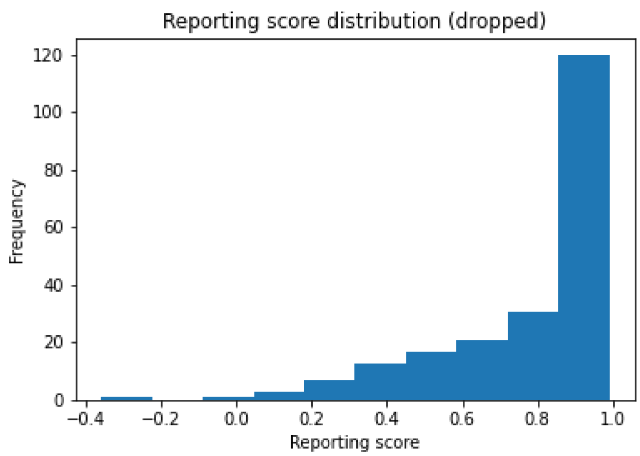
Appendix 8: Global relative reporting behavior. This plot shows the amount of cases reported per weekday as the proportion of cases reported for the whole week averaged over all countries. The error bars indicate a standard deviation of one.



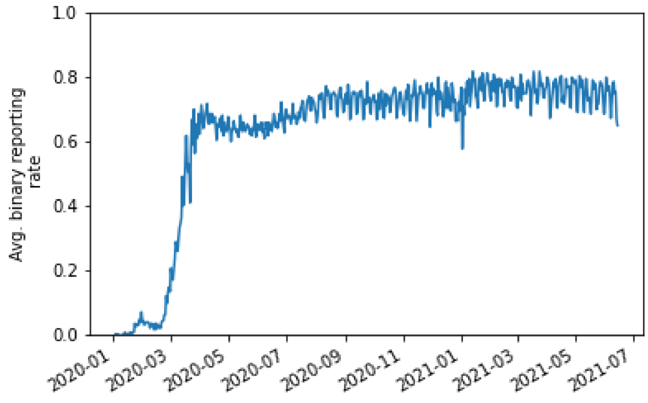
Appendix 9: Regional relative reporting behavior by weekday and WHO region. This plot shows the relative reporting behavior, i.e., the amount of cases reported per weekday as the proportion of cases reported for the whole week. The mean reporting rate over all countries for each WHO region is shown. The error bar around each dot indicate one standard deviation. The abbreviations stand for the African (AFRO), American (AMRO), Eastern Mediterranean (EMRO), European (EURO), South-East Asian (SEARO), and West Pacific (WPRO) WHO regional offices.



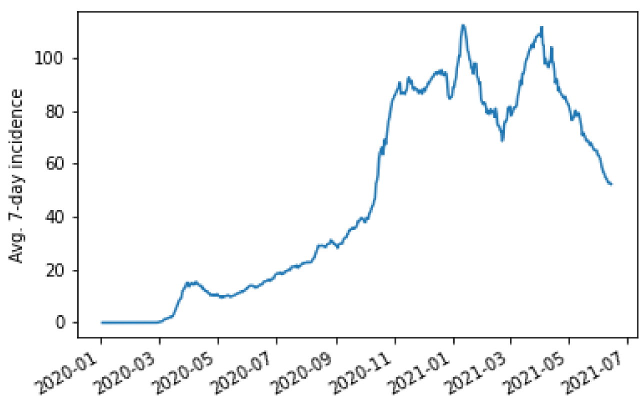
Appendix 10: Worldwide distribution of the reporting score. The reporting score is the difference of the binary reporting rate and the min-max scaled 7-day incidence rate. The 7-day incidence rate is the number of cases over the last 7 days per 100,000 inhabitants and the binary reporting rate quantifies the rate at which a country reported new cases COVID-19. This data was not imputed as opposed in the manuscript (c.f. Figure 1).



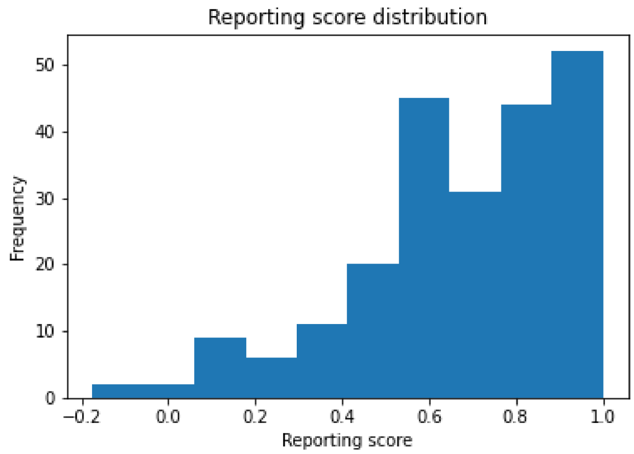
Appendix 11: Worldwide distribution of the reporting score. The reporting score is the difference of the binary reporting rate and the min-max scaled 7-day incidence rate. The 7-day incidence rate is the number of cases over the last 7 days per 100,000 inhabitants and the binary reporting rate quantifies the rate at which a country reported new cases COVID-19. Data was removed if a country reported less than seven times per week.



Appendix 12: The binary reporting rate averaged over all countries and over time. The binary reporting rate quantifies the rate at which a country reported new cases COVID-19.



Appendix 13: The 7-day incidence average over all countries and over time. The 7-day incidence rate is the number of cases over the last 7 days per 100,000 inhabitants.



Appendix 14: Worldwide distribution of the reporting score. Data from January to March 2020 was removed for a sensitivity analysis (c.f. Figure 1 of the manuscript). The reporting score is the difference of the binary reporting rate and the min-max scaled 7-day incidence rate. The 7-day incidence rate is the number of cases over the last 7 days per 100,000 inhabitants and the binary reporting rate quantifies the rate at which a country reported new cases COVID-19. Data was imputed if a country reported less than seven times per week.

Methods Appendix

To select three adequate time series, we constructed 14-day long time series of reported COVID-19 cases for all countries and over the entire time period. 14 days capture meaningful dynamics of infections while not being too long. Subsequently, we kept only time series in which more than zero cases were reported each day. We performed a linear regression on them and only kept time series with a positive slope since an increase of cases will require an *ad hoc* evaluation for mitigation measures opposed to a decrease. To select time series with low, medium and high number, we divided the remaining time series by quantile. Using the sum over the reported cases per time series, we calculated the 25%, 50%, and 75% quantile and picked the ones closest to those quantiles.

Using the three selected time series, we simulated delayed reporting. A delay was simulated by removing all reported cases from the delay period and adding them to the next successful reporting day. We looked at reporting delays from one to six days. For example, a country reported the following case numbers [10,14,22,15,7] from day one to day five. Assuming a two-day delay, the following reported cases [10,14,22,15,7] would turn into [0,0,24,0,0,37]. Cases from the first two-day delay period are reported on the third day (10+14=24). Cases from the third and fourth day are added to the fifth (22+15=37). The cases of the fifth day (7 cases) will appear on day seven. Finally, we calculated the relative difference of the real case counts and the ones caused by the delay per day and for each time series.