Syndromic surveillance for severe acute respiratory infections (SARI) enables valid estimation of COVID-19 hospitalization incidence and reveals underreporting of hospitalizations during pandemic peaks of three COVID-19 waves in Germany, 2020-2021

Authors
Kristin Tolksdorf¹, Walter Haas¹, Ekkehard Schuler², Lothar H. Wieler¹, Julia Schilling¹, Osamah Hamouda¹, Michaela Diercke¹, Silke Buda¹

Affiliations
1. Robert Koch Institute, Berlin, Germany
2. Helios Kliniken GmbH, Berlin, Germany

Abstract

Objective: With the emergence of coronavirus disease 2019 (COVID-19) countries had to establish COVID-19 surveillance by adapting existing systems, such as the mandatory notification system and syndromic surveillance systems. We estimated COVID-19 hospitalization and ICU burden from existing severe acute respiratory infections (SARI) surveillance and compared it to COVID-19 notification data.

Methods: Using data from the ICD-10 based hospital sentinel, we estimated SARI incidence and validated our estimations retrospectively using full population data. We estimated age-specific incidences for COVID-19 hospitalization and ICU for the first three COVID-19 waves in Germany, based on information of SARI cases (COVID-SARI). We compared these estimations to hospitalization and ICU burden of COVID-19 from notification data and described potential underreporting of hospitalizations in notification data.

Findings: The estimation of SARI incidence from sentinel data corresponded very well to full population data. The estimated COVID-SARI incidence matched the notified COVID-19 hospitalization incidence in the first wave, but was much higher during the second and the third wave. The proportion of unknown hospitalization status among notified COVID-19 cases was much higher in the later waves compared to the first wave. For all waves, the ICU incidence estimated from COVID-SARI was considerably higher than the results from notification data.

Conclusion: The use of SARI sentinel data adds valid and important information for assessing COVID-19 hospitalization and ICU burden, especially in times with high case numbers where notification data may be more incomplete.
Introduction

After the first cases of coronavirus disease 2019 (COVID-19) had been identified in Europe at the end of January, 2020, reporting systems had to be adapted to the newly emerging pandemic (1). The World Health Organization (WHO) supported countries worldwide with guidance on COVID-19 surveillance, indicating that regular revision would be necessary and country-specific adaptions may be needed (2, 3). In addition to active case finding and mandatory reporting of COVID-19 cases, the continued use of existing syndromic surveillance systems for respiratory diseases was recommended (4). Syndromic surveillance have been used to inform pandemic influenza severity assessments (PISA), providing information on transmissibility, seriousness of disease and impact on health systems (5). These indicators, developed in preparation for a potential new influenza pandemic, can be adapted for the COVID-19 pandemic as well. Considering the repeated emergence of new concerning variants of the SARS-CoV-2 virus and the waning of vaccine-acquired or natural immunity, the ongoing surveillance of severe COVID-19 cases is essential to assess disease seriousness (6, 7). However, the actual burden of hospitalization and even more of intensive care treatment in COVID-19 cases is difficult to determine, as valid data on hospitalized cases is neither easy to get nor easy to interpret.

In Germany, the mandatory reporting system for COVID-19 collects a wide range of epidemiological information on cases with laboratory confirmed SARS-CoV-2 infection, including information on hospitalization, intensive care and death of COVID-19 cases. However, information on hospitalization is not always possible to obtain, since continuous follow-up of cases by the local public health authorities is required. Prior analyses of notification data from the first COVID-19 wave has shown, that 13% of COVID-19 cases lacked information on hospitalization status (8). Among cases with known hospitalization status, about 10% of cases had no information on intensive care and about 50% did not have information on discharge of hospital. In Summer 2021, a new regulation was implemented to improve completeness of notification data regarding hospitalization. But especially during the peaks of the waves, lacking information on hospitalization is still expected. Thus, there is urgent need for additional, more thorough information on hospitalized patients to validate notification data on severe cases and to give a robust estimation of COVID-19 hospitalization and intensive care burden.

Here, we analyzed data from the ICD-10 based hospital sentinel for severe acute respiratory infections (ICOSARI) to estimate age-specific COVID-19 hospitalization and intensive care incidences, based on information of cases with severe acute respiratory infections (SARI) in accordance to WHO recommendations (2, 4, 5). Using data from three COVID-19 waves in Germany, we provide weekly estimates of severe COVID-19 burden for six age groups to show...
the validity of our method and point out the relevance of this important additional data source for a thorough understanding of the pandemic situation.

Methods

Data source

We use data from the syndromic hospital surveillance system ICOSARI (ICD-10 based hospital sentinel for severe acute respiratory infections) which was established in 2015 in cooperation with a large hospital network. The ICOSARI sentinel comprises 72 acute care hospitals in 13 of 16 federal states and covers about 6% of hospitalized patients in Germany. Within the participating hospitals, case-based data on admitted patients are routinely recorded. Data from the hospitals are collected and validated centrally within a data center. From there, anonymized validated case-based reports on all-cause admissions (denominator dataset) as well as reports including ICD-10 codes on all patients with any respiratory diagnosis (numerator dataset) are transferred weekly to the Robert Koch institute (RKI)(9). Weekly analyses are age-stratified and focused on SARI cases (ICD-10 primary diagnosis codes J09-J22, Basic Case Definition), which are extracted from the numerator dataset. The system has been validated and was described in detail in 2017 (10). Since influenza season 2016/17, the ICOSARI system is an integral component of the syndromic surveillance for acute respiratory infections, in combination with syndromic surveillance systems on population and on outpatient level (11, 12).

At the end of February, 2020, an ICD-10 diagnosis code for laboratory confirmed COVID-19 (ICD-10 secondary diagnosis code U07.1! (13, 14)) was introduced and implemented in the ICOSARI sentinel. Since March, 2020, we have enhanced the SARI monitoring to daily data transfers including preliminary ICD-10 codes of hospitalized patients. In particular, we have expanded our analyses to include the ICD-10 code for laboratory confirmed COVID-19 cases with SARI (COVID-SARI cases)(15).

From the mandatory reporting system for COVID-19, we used cases with known hospitalization (8). We included cases with hospitalization dates from week 10, 2020 to week 23, 2021 or, if no hospitalization date was reported, with reporting dates from week 10, 2020 to week 23, 2021.

Data from the ICOSARI sentinel and from the mandatory system was extracted on 29/11/2021.
Estimation of catchment population

Using the denominator dataset on all-cause admissions, we estimated the catchment population of the sentinel hospitals. We applied the proportional flow method as proposed by Norris and Bailey, using strata of German regions (each comprising more than one federal state), age groups and gender (16-18). Considering each stratum, we used the total number of admissions in the sentinel and the total number of patient admissions in Germany from full population data by the Federal Statistical Office to calculate the proportion of patient admissions within the sentinel (19). Subsequently, we derived the catchment population using the proportion of patient admissions and the total population in Germany for each stratum.

\[
\text{catchment population} = \frac{\text{admitted patients in sentinel}}{\text{admitted patients in Germany}} \times \text{population in Germany}
\]

We estimated the catchment population of the ICOSARI sentinel in the years 2015 to 2019. During the COVID-19 pandemic, the number of patient admissions (within the sentinel as well as country-wide) was largely affected by pandemic control and management measures. Thus, we used the median of the yearly estimations in 2015 to 2019 for each stratum to obtain the catchment population of the years 2020 and 2021.

We validated our method by comparing weekly SARI incidence estimations from the ICOSARI sentinel to the SARI incidence derived from full population data by the Federal Statistical Office using primary diagnoses of the years 2015 to 2019 (19).

Estimation of COVID-19 hospitalization and ICU treatment incidence

We estimated the incidence of COVID-SARI and of COVID-SARI with treatment in intermediate or intensive care units (COVID-SARI ICU), using primary and secondary diagnoses of cases with admission dates from week 10, 2020 to week 23, 2021 from the ICOSARI sentinel (data from 29/11/2021)(20-22). We used these estimations to evaluate hospitalization and intensive care burden of COVID-19 from notification data during three COVID-19 waves (data extracted on 29/11/2021). We assessed and described potential underreporting of hospitalization using the proportion of notified cases with unknown hospitalization status (23). We compared estimated hospitalization and intensive care incidences in 6 different age groups that have been established within the ICOSARI hospital sentinel (0 to 4 years, 5 to 14 years, 15 to 34 years, 35 to 59 years, 60 to 79 years, 80 years and above).
Results

Estimation of catchment population

In the years 2015 to 2019, the yearly number of patient admissions within the ICOSARI hospital sentinel ranged between 1.15 million and 1.18 million. The number of admissions was markedly lower in 2020. We thus estimated a total sentinel catchment population of 4.42 to 4.48 million for the years 2015 to 2019 and obtained a total catchment population of 4.46 million which we used as denominator for the years 2020 and 2021 (see table 1).

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of patient admissions in sentinel</th>
<th>Estimated catchment population in sentinel</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>1,109,192</td>
<td>4,418,814</td>
</tr>
<tr>
<td>2016</td>
<td>1,118,455</td>
<td>4,416,133</td>
</tr>
<tr>
<td>2017</td>
<td>1,111,084</td>
<td>4,417,856</td>
</tr>
<tr>
<td>2018</td>
<td>1,110,701</td>
<td>4,462,938</td>
</tr>
<tr>
<td>2019</td>
<td>1,112,756</td>
<td>4,477,524</td>
</tr>
<tr>
<td>2020</td>
<td>962,496</td>
<td>4,457,222</td>
</tr>
<tr>
<td>2021</td>
<td>851,112*</td>
<td>4,457,222</td>
</tr>
</tbody>
</table>

Table 1: Number of yearly patient admission in the ICOSARI sentinel and estimated catchment population (years 2015 to 2019: proportional flow, 2020 and 2021: stratified median.

*number of patient admissions for 2021 as of 29/11/2021

Figure 1 shows the weekly estimated SARI incidence per 100,000 population in the ICOSARI sentinel compared to the incidence retrospectively derived from the full population data by the Federal Statistical Office. Using the sentinel data, we overestimated the weekly SARI incidence by less than 1 SARI case per 100,000 population (Median [IQR]: 0.95 [0.42;1.65]), with larger differences in the peak weeks of the seasonal influenza epidemic (Fig. 1). Overall, the overlap of sentinel data and the full population data was very good regarding the timing and the level of trends and peaks.
Figure 1: Estimated incidence of SARI cases (ICD-10 codes J09 to J22), using ICOSARI sentinel data and full population data from the Federal Statistical Office by admission date.

**Estimation of COVID-19 hospitalization and ICU treatment incidence**

From week 10, 2020, to week 23, 2021, there were 19,530 COVID-SARI cases and 6,417 COVID-SARI ICU cases in the ICOSARI sentinel. Notification data comprised 260,745 hospitalized COVID-19 cases and 22,654 cases receiving intensive care. During the first wave (peak 1 - week 13, 2020: 6.6 vs 6.9 per 100,000 population), the estimated COVID-SARI incidence matched the COVID-19 hospitalization incidence from notification data very well. However, during the second and third wave (peak 2 - week 52, 2020: 22 vs 13 per 100,000 population; peak 3 - week 14, 2021: 14 vs. 8.9 per 100,000) (Fig. 3), the COVID-SARI incidence assumed higher values than the COVID-19 hospitalization incidence from notification data. Within the notification data, the median proportion of COVID-19 cases with unknown hospitalization status was 12% in the first wave (range: 11% to 14%). In the later waves however, the proportion was much higher with 23% (range: 14% to 36%) during the second wave and 18% (range: 16% to 21%) during the third wave (Fig. 2).

The estimated incidence of COVID-SARI ICU was much higher than the estimated incidence of COVID-19 ICU admissions derived from notification data. While COVID-SARI ICU incidence peaked at values 2.4, 7.3 and 4.7 per 100,000 in the three waves, the COVID-19 ICU incidence from notification data remained below 1 per 100,000 during the three peaks (0.86, 0.98 and 0.93 per 100,000 population) (Fig. 2). This considerable difference in ICU incidence estimations is also reflected by the proportion of ICU admissions among COVID-19 hospitalizations, which was 40%, 34% and 35% for the COVID-SARI patients in the sentinel, but only 15%, 7.4% and 11% in the notification data at the peaks 1, 2 and 3.
Figure 2: Estimated hospitalization incidence and intensive care incidence: COVID-19 cases with diagnosis code for SARI (ICD-10 codes J09-J22) from the hospital sentinel ICOSARI (solid line), for hospitalised COVID-19 cases from notification data (dashed line); proportion of COVID-19 cases with unknown hospitalization status from notification data (dotted line)

The estimations of hospitalization and intensive care incidence by age groups are shown in Figures 3 and 4. The incidence levels were very low in the age groups below 15 years of age during all three waves. In the age groups 35 years or older, the match in incidence estimations from the two data sources was very good in the first wave, but less in the later waves. Both data sources revealed the highest incidences in the age group 80 years and above during the first and the second wave. In the third wave, the hospitalization incidence in this age group was only slightly higher than the incidence of age group 60 to 79 years. Moreover, the ICU incidence of the age group 60 to 79 years was higher in the third wave than the ICU incidence of the oldest age group. Hospitalization and ICU incidence for the age group 35 to 59 years were lower during the second wave than during the third wave for both data sources.
Figure 3: Estimated hospitalization incidence by age groups: COVID-19 cases with diagnosis code for SARI (ICD-10 codes J09-J22) from the hospital sentinel ICOSARI (solid line) and for hospitalized COVID-19 cases from notification data (dashed line).

Figure 4: Estimated intensive care incidence by age groups: COVID-19 cases diagnosis code for SARI (ICD-10 codes J09-J22) from the hospital sentinel ICOSARI (solid line) and for hospitalised COVID-19 cases from notification data (dashed line).
Discussion

We estimated weekly hospitalization and intensive care incidences using syndromic hospital data of SARI cases. The burden of hospitalized COVID-19 cases was much larger in the second and in the third wave compared to the first wave. However, the extent of hospitalization burden was considerably larger when using the data on SARI cases compared to the notification data. Moreover, the intensive care incidence was remarkably lower in the notification data. Considering the high proportion of unknown hospitalization status during the second and the third wave, we assume underreporting of COVID-19 hospitalizations during the second and third wave and of intensive care treatments in general in the notification data.

Our analysis shows that the hospital sentinel provides more complete data on COVID-SARI in times when the pressure on the health system is very high, as was observed during the second and the third COVID-19 wave. In comparison to other European countries, the estimated hospitalization rates of COVID-SARI were on a similar level during the later waves, although the first wave was weaker in Germany (24-28). Several other countries in Europe and also on other continents had waves with very high case numbers during winter 2020/2021 followed by a wave during spring 2021. Remarkably, hospitalization and intensive care rates were lower overall and especially in the older age groups, whereas younger age groups tended to have even higher rates of severe cases in spring (24, 26-30). The reduced burden in the oldest age group is consistent with the start of the immunization program at the end of 2020 which mostly prioritized older age groups and effectively prevented severe cases in the most vulnerable population (31).

The rate of ICU among COVID-SARI patients in the ICOSARI sentinel is representative for German hospitalized COVID-19 patients and aligns with other sources for the first wave in Germany as well as in other countries (32-35). The gap in the intensive care incidence estimated from notification data and from sentinel data can largely be attributed to underreporting in the notification data. To some extent however, the definition of COVID-SARI may contribute to a higher ICU rate, as it includes only cases with severe respiratory infections and excludes COVID-19 cases without respiratory symptoms that were detected via admission screening in hospitals. However, both data sources consistently showed plausible trends regarding the differences in burden between age groups.

The strength of the hospital sentinel system is the daily electronic submission of quality-assured case-based data from hospital to RKI. The utilization of ICD-10 data is a digital standardized method without additional workload for the hospitals. It is both timely and enables international comparability. Thus, the system can give valid and complete data in times of high pressure on the health system, where local health authorities may be challenged with follow-up of COVID-19 cases, which is time and resource intensive. Moreover, the syndromic
surveillance system provides several years of SARI time series, allowing comparisons of COVID-SARI burden with the SARI burden during influenza waves (15, 36, 37).

However, our findings are subject to some limitations. The estimation of regional catchment population with the given data is difficult, therefore larger regions based on federal states were used (19). Yet, regional estimations are a strength of notification data.

For the years 2020 and 2021, we saw a notable decrease of patient admissions in the hospital sentinel. This effect was observed world-wide and was probably due to a combination of aspects such as the cancelling of elective operations and an increased hesitancy to use the health service in general and especially hospitals (38-45). As we had stable catchment population estimations in the preceding years, using a fixed catchment population derived from the median of previous years was justified.

The number of COVID-19-cases detected in hospitals is dependent on testing, and is likely underestimated. However, routine screening upon admission in sentinel hospitals was in place since July 2020. Furthermore, cases with laboratory confirmed SARS-CoV-2 infection hospitalized due to symptoms other than severe acute respiratory infections are not reported within the SARI sentinel. Thus, the focus on COVID-19 cases with SARI is more robust and less biased by the testing strategy as seen in notification data, as pneumonia was known to be the main syndrome of severe COVID-19 since the first case reports from China (46). We also note that weekly estimation of age-specific rates can lack accuracy due to low case numbers, especially in the younger age groups.

Conclusion

A valid link between the total number of notified COVID-19 cases, hospitalized patients and ICU patients is essential for the assessment of the seriousness of COVID-19, particularly considering the repeated emerging of new concerning variants (36, 47). Using data from the syndromic surveillance system ICOSARI allows valid estimations of hospitalization incidence, adds important information on intensive care burden in COVID-19 patients and reflects the different severity of the three waves of COVID-19 in Germany very well. Especially in times of high caseloads, the investigation of cases by local health authorities may lead to incomplete data on the severity of cases. In contrast, at that period the SARI sentinel provided robust estimations on the proportion of intensive care and on intensive care incidence. Thus, SARI surveillance is an important instrument to assess disease seriousness and to inform introduction or loosening of restrictions, additionally to the monitoring of transmission and of capacity utilisation (2, 25, 48, 49). Importantly, the use of an existing system has many
advantages such as known stability and potential biases, baseline data from previous years as well as known catchment population which is essential for the estimation of burden.

The method for the estimation of hospitalization and intensive care incidence from sentinel data can be applied not only in the COVID-19 pandemic. In particular, continuous surveillance of severe cases is essential to accompany the transition from a pandemic to an endemic disease, as well as in the SARI surveillance of seasonal influenza and RSV disease (50). Moreover, the surveillance and estimation method can be further adapted to other respiratory case definitions if needed. To improve regional representativeness, stepwise expansion of the system is planned.

References