

Surveillance of severe acute respiratory infections associated with SARS-CoV-2, influenza virus and RSV using ICD-10 codes: a case definition accuracy study across five European countries, 2021 to 2023

Miguel Angel Sanchez Ruiz¹, Diogo FP Marques¹, Frederikke Kristensen Lomholt², Lasse Skafte Vestergaard², Susana Monge^{3,4}, Marcos Lozano Álvarez^{3,5}, Gudrun Aspelund⁶, Marianna Thordardottir⁶, Ausra Dziugyte⁷, John-Paul Cauchi^{7,8}, Tjarda M Boere⁹, Irene K Veldhuijzen⁹, Elina Seppälä¹⁰, Håkon Bøås¹⁰, Trine Hesvevik Paulsen¹⁰, ESURE SARI group¹¹, Ausenda Machado¹², Ana Paula Rodrigues¹², Mariette Hooiveld¹³, Luis Alves de Sousa¹⁴, Ana Torres¹⁴, Carlos Carvalho¹⁴, Baltazar Nunes¹

1. Epiconcept, Paris, France
2. Department of Infectious Disease Epidemiology and Prevention, Statens Serum Institut, Copenhagen, Denmark
3. Department of Communicable Diseases, National Centre of Epidemiology, Institute of Health Carlos III, Madrid, Spain
4. Consortium for Biomedical Research on Infectious Diseases (CIBERINFEC), Madrid, Spain
5. Consortium for Biomedical Research on Epidemiology and Public Health (CIBERESP), Madrid, Spain.
6. Centre for Health Security and Communicable Disease Control, Directorate of Health, Reykjavik, Iceland
7. Infectious Disease Prevention and Control Unit (IDCU), Health Promotion and Disease Prevention, Msida, Malta
8. Faculty of Medicine & Dentistry (Malta Campus), Queen Mary University of London, VCT 2520, Victoria, Gozo, Malta
9. Centre for Infectious Disease Control, National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands
10. Department of Infection Control and Vaccines, Norwegian Institute of Public Health, Oslo, Norway
11. The members of the ESURE SARI group are listed under Collaborators
12. Epidemiology Department, National Institute of Health Doctor Ricardo Jorge, Lisbon, Portugal
13. Nivel, Utrecht, the Netherlands
14. European Centre for Disease Prevention and Control, Stockholm, Sweden

Correspondence: Miguel Angel Sanchez Ruiz (sanchez.r.miguel.angel@gmail.com)

Collaborators: The collaborators are listed at the end of the article.

Citation style for this article:

Sanchez-Ruiz Miguel-Angel, Marques Diogo FP, Lomholt Frederikke Kristensen, Vestergaard Lasse Skafte, Monge Susana, Lozano Álvarez Marcos, Aspelund Gudrun, Thordardottir Marianna, Dziugyte Ausra, Cauchi John-Paul, Boere Tjarda M, Veldhuijzen Irene K, Seppälä Elina, Bøås Håkon, Paulsen Trine Hesvevik, ESURE SARI group, Machado Ausenda, Rodrigues Ana Paula, Hooiveld Mariette, Alves de Sousa Luis, Torres Ana, Carvalho Carlos, Nunes Baltazar. Surveillance of severe acute respiratory infections associated with SARS-CoV-2, influenza virus and RSV using ICD-10 codes: a case definition accuracy study across five European countries, 2021 to 2023. *Euro Surveill.* 2025;30(27):pii=2400748. <https://doi.org/10.2807/1560-7917.ES.2025.30.27.2400748>.

Article received on 08 Nov 2024 / Accepted on 26 Feb 2025 / Published on 10 Jul 2025

Background: Surveillance of severe acute respiratory infections (SARI) using ICD-10 codes from electronic health records (EHR) lacks consensus on optimal case-defining codes. **Aim:** We determined codes that maximise sensitivity (Se) and positive predictive value (PPV) for SARI associated with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), influenza virus and respiratory syncytial virus (RSV) in Denmark, Iceland, Malta, Norway and Spain. **Methods:** We included hospitalisations from week 21/2021 to 39/2023, with ICD-10 diagnostic codes for respiratory disease (three-character codes J00–J99) or COVID-19 (U07.1, U07.2, country-specific codes for Denmark). We assessed Se and PPV of individual codes against laboratory results. Based on Se and PPV rank-sum, we selected the top 10 codes and combined them into 10 sets per pathogen. We identified sets that maximised the clinical utility index (CUI=Se × PPV), categorised as excellent (≥ 0.81), good (0.64–0.80), satisfactory (0.49–0.63) and poor (< 0.49). **Results:** We assessed 395,163 hospitalisations for SARI-SARS-CoV-2, 313,418 for SARI-influenza and 192,936 for SARI-RSV, all tested. For SARI-SARS-CoV-2, code U07.1 (B34.2A, B97.2A for Denmark) had excellent utility in Denmark,

Malta, Norway, Spain (≥ 0.82), and good utility in Iceland (0.79). For SARI-influenza, J09, J10 and J11 performed excellently in Denmark, Norway, Spain (≥ 0.83), satisfactorily in Malta (0.52), and poorly in Iceland (0.43). For SARI-RSV, J12, J20 and J21 achieved highest CUI but had poor utility (0.17–0.34). **Conclusions:** COVID-19- and influenza-specific three-character ICD-10 codes accurately identified SARI associated with SARS-CoV-2 and influenza virus. For SARI-RSV, four-character codes should be explored. We recommend context-specific assessments in countries adopting EHR-based surveillance.

Introduction

Among pathogens causing severe acute respiratory infections (SARI), severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), influenza and respiratory syncytial virus (RSV) have the largest impact in terms of morbidity, mortality, and absenteeism from work or education [1].

The surveillance of SARI enables the simultaneous monitoring of several respiratory pathogens using case definitions that traditionally rely on symptoms.

KEY PUBLIC HEALTH MESSAGE

What did you want to address in this study and why?

European countries are developing surveillance systems based on electronic health records (EHR) originally collected for clinical or administrative use. These records typically include diagnostic codes from the ICD-10 classification. We intended to find out which ICD-10 codes can be used to identify severe acute respiratory infections (SARI) associated with high-impact pathogens: SARS-CoV-2, influenza virus and respiratory syncytial virus (RSV).

What have we learnt from this study?

Across the participating countries (Denmark, Iceland, Malta, Norway, Spain), we found that similar diagnostic codes were the most suitable for identifying SARI associated with the three pathogens. The identified codes effectively detected SARI associated with SARS-CoV-2 and influenza virus, but were less successful in identifying RSV, highlighting the need to further explore codes for this pathogen.

What are the implications of your findings for public health?

The identified diagnostic codes can be used in the participating countries to monitor trends or understand the burden of SARI associated with specific pathogens, thereby informing public health decisions. We applied a common methodology to identify ICD-10 codes for detecting SARI associated with SARS-CoV-2, influenza virus and RSV. This can be replicated by other countries currently implementing EHR-based SARI surveillance in Europe.

The World Health Organization (WHO) defines a SARI as an acute respiratory infection with a history of fever or measured temperature of $\geq 38^{\circ}\text{C}$, and cough, with onset within the last 10 days, that requires hospitalisation [2].

Since 2020, most countries in the European Region have established SARI surveillance systems which exhibit considerable heterogeneity [3-14]. Some countries rely on electronic health records (EHR) [9-13] collected and stored for clinical or reimbursement purposes, which may include diagnosis codes such as the 10th edition of the International Classification of Disease codes (ICD-10) [15]. When symptom data are unavailable or not collected in a standardised manner, ICD-10 diagnosis codes may serve as a proxy for identifying SARI associated with specific pathogens [9-13]. However, more studies are required to determine the most suitable ICD-10 codes that define SARI for surveillance purposes and to assess whether common definitions can be used across countries. The absence of this information hampers the implementation of EHR-based SARI surveillance, which complicates the estimation, comparison and pooling of national incidence rates across Europe.

We conducted a study to determine ICD-10 code-based case definitions that maximise the sensitivity (Se) and positive predictive value (PPV) in identifying SARI associated with one specific pathogen (SARS-CoV-2, influenza virus or RSV) and SARI associated with any of the

three. This was assessed overall, by age group, and considering national-level SARI activity.

Methods

Study design

The study was a retrospective assessment of the accuracy of ICD-10 code-based case definitions to identify SARI associated with SARS-CoV-2, influenza virus or RSV. The reference standard was a pathogen-specific laboratory test result, primarily determined by PCR, with rapid antigen testing (RAT) used alone in 10% or fewer of the samples.

Study population and study setting

Data were extracted from SARI surveillance systems in five countries and analysed using a common protocol. Data items included the week of hospital admission, ICD-10 diagnosis codes, age group, and pathogen-specific laboratory testing. We did not disaggregate by sex as this is not expected to enhance case identification. Denmark, Iceland, Malta and Norway used nationwide EHR-based surveillance databases with nearly complete coverage of their populations. In contrast, Spain used data from sentinel hospitals in four regions (the Basque Country, Catalonia, the Valencian Community and Galicia) representing 8.3% of the national population.

Inclusion criteria

We selected records through a two-step process. Firstly, countries identified hospitalisation episodes

TABLE 1

Study population according to SARI surveillance systems, Denmark, Iceland, Malta, Norway, Spain, week 21/2021–39/2023

Country	Patients included	Definitions		ICD-10 codes		Laboratory considerations	
		Hospital admission	Readmission	Accessible	Code class ^a	Pathogens included and type of test	Time window for laboratory confirmation
Denmark	All hospitalisation episodes meeting inclusion criteria	Hospitalised for at least 12 h	Subsequent hospital stays within 14 days from discharge	J00–J99+country-specific SARS-CoV-2 codes ^b	Discharge	SARS-CoV-2, influenza virus, RSV (PCR)	From 10 days before admission to 3 days after admission
Iceland	All hospitalisation episodes meeting inclusion criteria	Admission does not depend on length, can be hours or a day	Subsequent hospital stays within 2 days from discharge	J00–J99+U07.1, U07.2	Admission and discharge	SARS-CoV-2, influenza (PCR)	From 10 days before admission to 2 days after admission
Malta	All hospitalisation episodes meeting inclusion criteria	Hospitalised for more than 24 h	No use of readmissions, any new contact with a hospital is considered a new admission	J00–J99+U07.1, U07.2	Discharge	SARS-CoV-2 (RAT and PCR); influenza virus and RSV (PCR)	From 14 days before admission to 3 days after (3 days is a proxy to 48 h within hospital)
Norway	All hospitalisation episodes meeting inclusion criteria	Hospital admissions are identified by a specific variable that categorises the patient contacts as polyclinical, day treatment, or admission	Subsequent hospital stays with no more than 1 day in between	J00–J22, J80+U07.1, U07.2, J30–J99 excluding J80 are grouped and referred to as J-non-SARI	Discharge	SARS-CoV-2 (RAT and PCR), influenza virus and RSV (PCR)	From 14 days before admission to 2 days after discharge
Spain	Hospitalisation episodes selected as per the national SARI sentinel surveillance protocol and meeting the study inclusion criteria ^c	Hospitalised for more than 24 h, exclusions ^d	Subsequent hospital stays within 14 days from discharge	J00–J99+U07	Discharge	SARS-CoV-2, influenza virus and RSV (RAT and PCR)	Sample collection is recommended as soon as possible after admission, and preferably within 7 days of symptom onset

ICD-10: 10th edition of the International Classification of Disease codes; RAT: rapid antigen test; RSV: respiratory syncytial virus; SARI: severe acute respiratory infection; SARS-CoV-2: severe acute respiratory syndrome coronavirus 2.

^a Both principal and secondary ICD-10 codes were included in all countries. In Denmark, the surveillance system does not classify codes as either admission or discharge, instead, codes are updated as information becomes available. Since this was a retrospective study, the extracted codes are expected to correspond to discharge ones mostly. In Norway, diagnostic codes are typically registered at discharge, however, in some cases, codes may be recorded as soon as the diagnosis is confirmed, potentially before discharge. Malta and Spain used discharge codes. Iceland used a mix of admission and discharge codes, but these could not be distinguished from one another.

^b Denmark used country-specific codes for COVID-19, as follows: B97.2A2A: COVID-19 severe acute respiratory syndrome B34.2A2A: COVID-19, not elsewhere classified, Z038PA1: observation due to suspicion of COVID-19 infection.

^c As per the Spanish sentinel SARI surveillance protocol, all participating regions but the Basque Country pre-selected admissions compatible with SARI (according to locally defined clinical criteria or at the clinician's discretion) from specific weekdays (1 day or 2 days a week) and subsequently applied the study inclusion criteria (i.e. hospitalisation episodes having a respiratory disease or COVID-19-specific ICD-10 code were retained for the study). In the Basque Country, the study inclusion criteria were directly applied to all patients hospitalised.

^d Exclusions applied in Spain for the definition of hospital admissions: patients under observation in the emergency room, day hospital, haemodialysis sessions, outpatient surgery, or transfers between the hospital's services are not considered hospital admissions.

from patients of all ages who had been routinely assigned at least one three-character ICD-10 diagnosis code for respiratory disease (J00 to J99) or a four-character COVID-19 ICD-10 code (U07.1, U07.2, Denmark used country-specific codes listed in the Supplement). Countries used principal and secondary diagnosis codes assigned mostly at discharge. In Iceland, while the database included both admission and discharge codes, the available information did not allow determining whether a code corresponded to admission or discharge diagnoses. Countries used ICD-10 versions in their national languages. Secondly, we retained those of the identified hospitalisation episodes in which patients had been tested for SARS-CoV-2, influenza

viruses (A or B) and/or RSV. The use of PCR (primarily) or RAT, and the time window for laboratory results to be linked to a hospitalisation episode, varied according to country-specific surveillance standards. The details of the inclusion process and definitions per country are described in Table 1. Supplementary Table S1 outlines the testing practices.

Study period

We included hospitalisation episodes between week 21/2021 and week 39/2023. In Malta, patients younger than 18 years were only included in the SARI surveillance system from week 6/2023. National surveillance teams extracted data in or after December 2023.

TABLE 2

Two-by-two table for assessing the accuracy of single-code and code combinations against a laboratory-defined reference standard, SARI surveillance, Denmark, Iceland, Malta, Norway, Spain, week 21/2021–39/2023

ICD-10 code(s) under assessment	Reference standard		Total
	Hospitalisations positive for pathogen p (p) (pathogen present)	Hospitalisations negative for pathogen p (p') (pathogen absent)	
Hospitalisations with ICD-10 code(s) under assessment (case definition a (a) met)	n_{ap}	$n_{ap'}$	$n_{a.} = n_{ap} + n_{ap'}$
Hospitalisations without ICD-10 code(s) under assessment (case definition a (a') unmet)	$n_{a'p}$	$n_{a'p'}$	$n_{a'.} = n_{a'p} + n_{a'p'}$
Total	$n_{.p} = n_{ap} + n_{a'p}$	$n_{.p'} = n_{ap'} + n_{a'p'}$	$n_{a.} + n_{a'.} = n_{.p} + n_{.p'}$

a' and p' represent the complementary of a and p , respectively. n_{ap} : true positive cases, $n_{ap'}$: false positive cases, $n_{a'p}$: false negative cases, $n_{a'p'}$: true negative cases; $n_{a.}$: sum of n_{ap} and $n_{ap'}$; $n_{a'.}$: sum of $n_{a'p}$ and $n_{a'p'}$; $n_{.p}$: sum of n_{ap} and $n_{a'p}$; $n_{.p'}$: sum of $n_{ap'}$ and $n_{a'p'}$.

Data analysis

Descriptive analysis

We described the study population and testing patterns according to age group and national-level SARI activity (i.e. high/low weekly SARI notification rates; for a full list of practices by country see Supplementary Table S2). We described testing patterns as the proportion of hospitalisation episodes in which patients were tested for SARS-CoV-2, influenza virus and/or RSV, along with the proportion of pathogen-specific positive tests.

Accuracy assessment

The reference standard for SARI associated with a specific pathogen included hospitalisation episodes from patients with respiratory disease or COVID-19 diagnosis codes and laboratory results for that pathogen (primarily PCR, but also RATs). Laboratory confirmation defined true positive pathogen-specific SARI. For SARI associated with any considered pathogens, true positives were defined by laboratory confirmation of SARS-CoV-2, influenza or RSV in patients tested for all three pathogens during the same hospitalisation. All available codes per patient, whether assigned at discharge or admission (the latter available only in Iceland), corresponding to principal or secondary diagnosis, were included in the analysis. This approach aimed to enhance case definition sensitivity and ensure the identification of patients who may not have been assigned a respiratory disease or COVID-19 code as their principal diagnosis.

We assessed the accuracy of the various code sets or case definitions following a three-step process:

-1- Building case definitions based on single-code accuracy

We calculated the Se and PPV of each ICD-10 code (J00-J99, U07.1, U07.2, Danish COVID-19 codes) for identifying a specific pathogen among SARS-CoV-2, influenza virus and RSV, or any of the three. The Se of

an ICD-10 code a for identifying hospitalisations with laboratory confirmation of pathogen p (p could equal SARS-CoV-2, influenza virus, RSV or any of the three), was given by

$$Se_{ap} = \frac{n_{ap}}{n_{.p}}$$

Here, n_{ap} represents the number of true positive cases (TP), i.e. hospitalisation episodes from patients with the relevant ICD-10 code a and laboratory confirmation of pathogen p , and $n_{.p}$ denotes the total number of hospitalisations with laboratory confirmation of pathogen p . The PPV of an ICD-10 code a for identifying hospitalisations positive for pathogen p was given by

$$PPV_{ap} = \frac{n_{ap}}{n_{a.}}$$

Here, $n_{a.}$ refers to all hospitalisations for which a specific ICD-10 code a had been assigned, regardless of their laboratory results. Both Se and PPV were expressed as percentages (Table 2).

For each single ICD-10 code a , we assigned Se and PPV ranks for the identification of pathogen p . Codes with higher Se received lower Se ranks $R(Se_{ap})$, and codes with higher PPV received lower PPV ranks $R(PPV_{ap})$, and vice versa. Then, for each ICD-10 code a , we calculated a score based on the rank sums of Se and PPV to identify pathogen p , defined as

$$Score_{ap} = R(Se_{ap}) + R(PPV_{ap})$$

We considered that the lower the $Score_{ap}$, the higher the ability of a code to maximise Se and PPV for identifying

TABLE 3

Testing status and testing results for hospitalisations with respiratory disease and COVID-19 ICD-10 codes, Denmark, Iceland, Malta, Norway, Spain, week 21/2021–week 39/2023

Country	Hospitalisations meeting inclusion criteria ^a	n tested for at least one pathogen ^b	SARS-CoV-2			Influenza virus			RSV			SARS-CoV-2, influenza, RSV					
			Tested (among hospitalised)		Test-positive (among tested)	Tested (among hospitalised)		Test-positive (among tested)	Tested (among hospitalised)		Test-positive (among tested)	Tested for all three (among hospitalised)		Test-positive for at least one (among tested)			
			n	%	n	%	n	%	n	%	n	%	n	%	n	%	
Denmark	237,208	179,051	74.3	32,832	18.6	134,842	56.8	6,522	4.8	71,563	30.2	6,680	9.3	68,594	28.9	19,237	28.0
Iceland	7,052	4,428	60.5	1,632	38.2	1,834	26.0	179	9.8	NA	NA	NA	NA	NA	NA	NA	NA
Malta	14,066	11,967	67.6	2,226	23.4	6,950	49.4	403	5.8	6,950	49.4	321	4.6	4,486	31.9	711	15.9
Norway	271,097	177,884	65.1	37,918	21.5	159,354	58.8	8,908	5.6	112,802	41.6	8,453	7.5	111,687	41.2	34,301	30.7
Spain	35,161	28,976	81.0	6,435	22.6	10,438	29.7	1,677	16.1	1,621	4.6	280	17.3	1,584	4.5	649	41.0
Total	564,584	402,306	70.0	81,043	20.5	313,418	55.5	17,689	5.6	192,936	34.2	15,734	8.2	186,351	33.0	54,898	29.5

ICD-10: 10th edition of the International Classification of Disease codes; NA: not available; RSV: respiratory syncytial virus; SARS-CoV-2: severe acute respiratory syndrome coronavirus 2.

^a Hospitalisations with respiratory disease ICD-10 codes: J00–J99 or COVID-19 ICD-10 codes: U07.1, U07.2, and Danish codes B34.2A, B97.2A, Z03.8PA1.

^b Number tested for at least one of SARS-CoV-2, influenza virus or RSV.

TABLE 4

Accuracy of the ICD-10 code-based case definitions that maximise the sensitivity and positive predictive value in identifying SARI associated with SARS-CoV-2, influenza virus, RSV or any of the three viruses, Denmark, Iceland, Malta, Norway, Spain, week 21/2021–week 39/2023

All ages	ICD-10 ^a							TP+FN	TP+FP	Se (%) ^b	PPV (%) ^c	CUI ^d	CUI score ^e
	U07.1	J09	J10	J11	J12	J20	J21						
	B34.2A+B97.2A												
SARI-SARS-CoV-2													
Denmark	✓							32,832	31,027	90.2	95.4	0.86	E
Iceland	✓							1,632	1,943	96.7	81.2	0.79	G
Malta	✓				X			2,226	2,246	91.1	90.2	0.82	E
Norway	✓							37,918	33,973	88.4	98.7	0.87	E
Spain	✓							6,435	6,482	95.7	95.0	0.91	E
SARI-influenza													
Denmark		✓	✓	✓				6,522	5,612	84.5	98.1	0.83	E
Iceland		✓	✓	✓				179	214	71.5	59.8	0.43	P
Malta	X	X	✓	✓		X		403	240	55.6	93.3	0.52	S
Norway		✓	✓	✓				8,908	8,449	91.6	96.5	0.88	E
Spain		✓	✓	X				1,677	1,599	93.7	98.3	0.92	E
SARI-RSV													
Denmark					✓	✓	✓	6,680	6,805	58.9	57.8	0.34	P
Malta					X	X	✓	321	228	34.6	48.7	0.17	P
Norway					X	X	✓	8,453	4,550	38.2	71.1	0.27	P
Spain					X	X	✓	280	238	49.6	58.4	0.29	P
SARI-any of the three pathogens													
Denmark	✓	✓	✓	X	✓	✓	✓	19,237	18,105	79.5	84.5	0.67	G
Malta	✓	X	✓	✓	X	X	✓	711	637	69.8	77.9	0.54	S
Norway	✓	X	✓	✓	✓	X	✓	34,301	33,442	83.5	85.6	0.71	G
Spain	✓	X	✓	X	✓	X	✓	649	638	78.9	80.3	0.63	S

✓: The ICD-10 code maximises the sensitivity and positive predictive value in identifying SARI associated with the pathogen under assessment; X: The ICD-10 code does not maximise the sensitivity and positive predictive value in identifying SARI associated with the pathogen under assessment; CUI: clinical utility index; FN: false negative; FP: false positive; ICD-10: 10th edition of the International Classification of Disease; PPV: positive predictive value; RSV: respiratory syncytial virus; SARI: severe acute respiratory infection; SARS-CoV-2: severe acute respiratory syndrome coronavirus 2; Se: sensitivity; TP: true positive.

^a ICD-10 codes; U07.1: COVID-19, virus identified. B34.2A: COVID-19, not elsewhere classified (Denmark-specific code). B97.2A: COVID-19 severe acute respiratory syndrome (Denmark-specific code). J09: Influenza due to identified zoonotic or pandemic influenza virus. J10: Influenza due to identified seasonal influenza virus. J11: Influenza due to unidentified influenza virus. J12: Viral pneumonia, not elsewhere classified. J20: Acute bronchitis. J21: Acute bronchiolitis.

^b $Se(\%) = TP / (TP + FN) \times 100$.

^c $PPV(\%) = TP / (TP + FP) \times 100$.

^d $CUI = Se \times PPV$.

^e CUI score utilises the following qualitative categories; E: excellent utility; G: good utility; S: satisfactory utility; P: poor utility.

SARI-pathogen refers to SARI associated with a specific pathogen (SARS-CoV-2, influenza virus, RSV or any of the three).

pathogen p . If two codes achieved the same score, the code associated with the higher n_{ap} was prioritised.

We then determined the 10 ICD-10 codes with the lowest Score_{ap} for each pathogen p . These were sequentially combined according to their rank order to formulate 10 cumulative pathogen-specific SARI case definitions (SARI- p) consisting of one to 10 diagnosis codes. The first SARI- p case definition included the single ICD-10 code a with the lowest score, and the last SARI- p case definition included the 10 ICD-10 codes with the lowest scores.

-2- Case-definition accuracy assessment

We calculated the Se and PPV of the 10 ICD-10 code-based case definitions developed as described above to identify a specific pathogen or any of the three. In a

similar manner as with single codes, the Se and PPV of case definitions were given by

$$Se_{ap} = \frac{n_{ap}}{n_p} \text{ and } PPV_{ap} = \frac{n_{ap}}{n_a}$$

However, here n_{ap} or TP represent hospitalisation episodes containing at least one of the codes in the case definition a with laboratory confirmation, and n_a refers to all hospitalisations assigned at least one of the codes in the case definition a , regardless of their laboratory results (Table 2).

-3- Case-definition selection

We selected case definitions that maximised the clinical utility index ($CUI = Se \times PPV$). We categorised the

CUI based on the score as excellent (≥ 0.81), good (0.64–0.80), satisfactory (0.49–0.63) or poor (< 0.49) utility [16,17]. In instances where two case definitions achieved the same CUI, the one incorporating fewer codes was chosen.

We based our analysis on the Se and PPV due to the following considerations: the Se and specificity (Sp) are measures of accuracy concerning the occurrence of TP and true negative (TN) cases, respectively. The Se quantifies the capability of a case definition to detect a condition or pathogen when it is indeed present. Meanwhile, the PPV and the negative predictive value (NPV) are accuracy measures of discrimination that help ascertain the actual status (laboratory result) of, respectively, cases and non-cases classified by the case definition. The PPV provides the proportion of TP among cases identified by the case definitions. Therefore, case definitions with high Se and PPV enhance the detection of TP cases while reducing the number of FP and are preferable for case finding. Consequently, we considered that the higher the CUI, the higher the suitability of a code set for finding SARI cases associated with specific pathogens [16,18].

Other considerations

Denmark, Malta, Norway and Spain completed the above analysis. However, Iceland's data did not include RSV testing results, which limited their analysis to SARS-CoV-2 and influenza virus.

Stratification was performed to identify heterogeneity due to patient characteristics (age groups: 0–4, 5–14, 15–64 and ≥ 65 years), and high or low SARI activity. The period of high SARI activity included all weeks where the number of SARI admissions per 100,000 population exceeded the 66th percentile of the admissions within the study period, as defined by each country. Weeks with admissions below this threshold formed the low-activity period. A description of periods by SARI activity and country is appended in Supplementary Table S2.

Results

Study population and testing results

Study population

A total of 564,584 hospitalisation episodes met the ICD-10 code selection criteria and were therefore included. Norway and Denmark contributed the largest number of hospitalisations, representing 48.0% ($n=271,097$) and 42.0% ($n=237,208$) of the total, respectively. Spain, Malta and Iceland provided smaller proportions, contributing 6.2% ($n=35,161$), 2.5% ($n=14,066$) and 1.2% ($n=7,052$) of the hospitalisation episodes, respectively (Table 3).

Proportion tested

Among all 564,584 hospitalisations, 71.3% ($n=402,306$) had a test reported for at least one of the pathogens

under consideration, 70.0% ($n=395,163$) had a test reported for SARS-CoV-2, 55.5% ($n=313,418$) had a test for influenza virus, 34.2% ($n=192,936$) for RSV, and 33.0% ($n=186,351$) for all three pathogens. The proportion tested for SARS-CoV-2 ranged from 60.5% (Iceland) to 81.0% (Spain), while the proportion tested for influenza virus ranged from 26.0% (Iceland) to 58.8% (Norway), and for RSV from 4.6% (Spain) to 49.4% (Malta). The proportion tested for all three viruses ranged from 4.5% (Spain) to 41.2% (Norway) (Table 3). In Supplementary Table S3, we append additional information on the proportion tested and the test positivity by SARI activity and age group.

Test positivity

Among hospitalisations with reported laboratory tests, 20.5% (81,043/395,163), 5.6% (17,689/313,418) and 8.2% (15,734/192,936) tested positive for SARS-CoV-2, influenza virus, and RSV, respectively. For SARS-CoV-2, the test positivity ranged from 18.6% (Denmark) to 38.2% (Iceland), while for influenza virus the test positivity ranged from 4.8% (Denmark) to 16.1% (Spain), and for RSV from 4.6% (Malta) to 17.3% (Spain). Among hospitalisations from patients tested for all three pathogens, 29.5% (54,898/186,351) had laboratory confirmation for at least one of the pathogens, ranging from 15.9% (Malta) to 41.0% (Spain) (Table 3).

Case definition accuracy assessment

Tables 4 and 5 show overall and stratified accuracy results for case definitions that achieved the highest CUI, respectively. Supplementary Table S4 provides the ICD-10 codes selected to build pathogen-specific SARI case definitions, along with the Se, PPV and CUI of the 10 case definitions assessed for each pathogen and stratum.

SARI associated with SARS-CoV-2

The best accuracy for identifying SARI-SARS-CoV-2 cases was achieved by a case definition that included only U07.1 ('COVID-19, virus identified') in all countries, except Denmark, where B34.2A ('COVID-19, not elsewhere classified') and B97.2A ('COVID-19 severe acute respiratory syndrome') were the best. These codes yielded excellent utility across most countries and strata (Se $\geq 88\%$, PPV $\geq 90\%$, CUI ≥ 0.82 for the all-ages group in Denmark, Malta, Norway and Spain). Good utility was reported for U07.1 in the ≥ 65 -year-old group, during periods of high SARI activity, and the all-ages group in Iceland. In Denmark, B34.2A and B97.2A also achieved good utility for the 5–14-year-old group. In Spain, U07.1 yielded good and satisfactory utility for the 0–4-year-old and the 5–14-year-old groups, respectively.

SARI associated with influenza virus

Case definitions including codes J09 ('Influenza due to identified zoonotic or pandemic influenza virus'), J10 ('Influenza due to identified seasonal influenza virus'), and J11 ('Influenza, virus not identified') yielded the best accuracy results for SARI-influenza case finding

TABLE 5A

Accuracy of the ICD-10 code-based case definitions that maximise the sensitivity and positive predictive value in identifying SARI cases associated with SARS-CoV-2, influenza virus, RSV or any of the three pathogens, by SARI activity and age group, Denmark, Iceland, Malta, Norway, Spain, week 21/2021–week 39/2023

Country	SARI-SARS-CoV-2				SARI-influenza				SARI-RSV				SARI-any of the three pathogens							
	ICD-10 ^a	Se ^b (%)	PPV ^c (%)	CUI ^d	ICD-10 ^a	Se ^b (%)	PPV ^c (%)	CUI ^d	ICD-10 ^a	Se ^b (%)	PPV ^c (%)	CUI ^d	ICD-10 ^a	Se ^b (%)	PPV ^c (%)	CUI ^d				
Period of high SARI activity																				
Denmark	B34,B97	91.3	95.3	0.87	E	J09, J10	85.1	98.5	0.84	E	J12, J20, J21	55.5	61.0	0.34	P	B34, B97, J09, J10, J12, J20, J21	79.4	86.4	0.69	G
Iceland	U07.1	97.1	78.4	0.76	G	J09, J10, J11	77.0	60.5	0.47	P	NA				NA					
Malta	U07.1	92.4	91.3	0.84	E	J10, J11	60.5	94.9	0.57	S	J21	17.3	26.0	0.05	P	U07.1, J10, J11	70.5	86.6	0.61	S
Norway	U07.1	88.5	98.8	0.87	E	J09, J10, J11	92.4	97.6	0.90	E	J12, J20, J21	82.8	46.3	0.38	P	U07.1, J10, J12, J21	82.4	90.5	0.75	G
Spain	U07.1	96.5	95.0	0.92	E	J09, J10	94.7	97.8	0.93	E	J21	51.7	68.2	0.35	P	U07.1, J10, J12, J21	78.4	84.4	0.66	G
Period of low SARI activity																				
Denmark	B34,B97	88.5	95.6	0.85	E	J09, J10, J11	82.7	98.1	0.81	E	J12, J20, J21	63.4	54.6	0.35	P	B34, B97, J09, J12, J21	72.7	92.8	0.67	G
Iceland	U07.1	95.9	86.9	0.83	E	J09, J10, J11	54.5	57.1	0.31	P	NA				NA					
Malta	U07.1	90.7	90.0	0.82	E	J10, J11	53.4	92.5	0.49	S	J21	39.8	55.1	0.22	P	U07.1, J10, J11, J21	68.5	76.9	0.53	S
Norway	U07.1	88.4	98.6	0.87	E	J09, J10, J11	89.7	94.1	0.84	E	J12, J21	64.3	20.7	0.13	P	U07.1, J10, J12, J21	83.0	78.4	0.65	G
Spain	U07.1	94.7	95.1	0.90	E	J09, J10, J11	95.4	98.7	0.94	E	J21	39.6	30.6	0.12	P	U07.1, J10, J12	69.6	88.1	0.61	S
Age 0–4 years																				
Denmark	B34	82.9	97.8	0.81	E	J09, J10	74.5	95.2	0.71	G	J12, J20, J21, J96	81.3	55.4	0.45	P	B34, J09, J10, J12, J21	62.5	82.8	0.52	S
Iceland	U07.1	95.6	95.6	0.91	E	U07.1, J09, J10, J91	81.0	51.5	0.42	P	NA				NA					
Malta	U07.1	95.4	96.9	0.93	E	J10	46.2	80.0	0.37	P	J21	71.9	51.0	0.37	P	U07.1, J21	76.2	60.6	0.46	P
Norway	U07.1	86.4	98.6	0.85	E	J09, J10, J11	88.0	89.2	0.79	G	J12, J21	79.8	63.7	0.51	S	U07.1, J10, J21	74.2	80.8	0.60	S
Spain	U07.1	89.6	89.6	0.80	G	J10	91.0	98.4	0.90	E	J12, J21	67.1	60.4	0.41	P	U07.1, J10, J12, J21	72.6	67.1	0.49	P
Age 5–14 years																				
Denmark	B34,B97	86.4	92.5	0.80	G	J09, J10	78.5	95.2	0.75	G	J21	11.9	66.7	0.08	P	J09, B34, J10, B97	66.8	93.2	0.62	S
Iceland	U07.1	95.6	95.6	0.91	E	J03, J04, J10, J11	71.4	62.5	0.45	P	NA				NA					
Malta	U07.1	100	100	1.0	E	J10, J11	57.9	100.0	0.58	S	J06, J18	75.0	14.3	0.11	P	U07.1, J06, J10, J11, J21, J18	87.5	63.6	0.56	S
Norway	U07.1	84.5	98.0	0.83	E	J09, J10, J11	91.3	96.9	0.89	E	J12, J20, J21	61.0	24.1	0.15	P	U07.1, J10, J11	71.0	97.4	0.69	G
Spain	U07.1	83.3	71.4	0.60	S	J10	81.8	94.7	0.78	G	J12	50.0	50.0	0.25	P	U07.1, J10, J12, J95, J99	64.3	90.0	0.58	S

CUI: clinical utility index; FN: false negative; FP: false positive; ICD-10: 10th edition of the International Classification of Disease codes; PPV: positive predictive value; RSV: respiratory syncytial virus; SARI: severe acute respiratory infection; SARS-CoV-2: severe acute respiratory syndrome coronavirus 2; Se: sensitivity; TP: true positive.

^a ICD-10 codes; U07.1: COVID-19, virus identified. B34.2A: COVID-19, not elsewhere classified (Denmark-specific code). B97.2A: COVID-19 severe acute respiratory syndrome (Denmark-specific code). J03: Acute tonsillitis. J04: Acute laryngitis and tracheitis. J06: Acute upper respiratory infections of multiple and unspecified sites. J09: Influenza due to identified zoonotic or pandemic influenza virus. J10: Influenza due to identified seasonal influenza virus. J11: Influenza due to unidentified influenza viruses. J12: Viral pneumonia, not elsewhere classified. J14: Pneumonia due to *Haemophilus influenzae*. J15: Bacterial pneumonia, not elsewhere classified. J16: Chlamydial pneumonia. J18: Pneumonia, unspecified organism. J20: Acute bronchitis. J21: Acute bronchiolitis. J22: Unspecified acute lower respiratory infection. J40: Bronchitis, not specified as acute or chronic. J44: Other chronic obstructive pulmonary disease. J45: Asthma. J60: Coalworker's pneumoconiosis. Asthma. J91: Pleural effusion in conditions classified elsewhere. J95: Postprocedural respiratory disorders, not elsewhere classified. J96: Respiratory failure, not elsewhere classified. J98: Other respiratory disorders. J99: Respiratory disorders in diseases classified elsewhere.

^b Se (%) = TP / (TP + FN) × 100.

^c PPV (%) = TP / (TP + FP) × 100.

^d CUI = Se × PPV; results are provided both quantitatively and qualitatively, using the following categories: E: excellent utility, G: good utility, S: satisfactory utility, P: poor utility.

TABLE 5B

Accuracy of the ICD-10 code-based case definitions that maximise the sensitivity and positive predictive value in identifying SARI cases associated with SARS-CoV-2, influenza virus, RSV or any of the three pathogens, by SARI activity and age group, Denmark, Iceland, Malta, Norway, Spain, week 21/2021–week 39/2023

Country	SARI-SARS-CoV-2					SARI-influenza					SARI-RSV					SARI-any of the three pathogens					
	ICD-10 ^a	Se ^b (%)	PPV ^c (%)	CUI ^d		ICD-10 ^a	Se ^b (%)	PPV ^c (%)	CUI ^d		ICD-10 ^a	Se ^b (%)	PPV ^c (%)	CUI ^d		ICD-10 ^a	Se ^b (%)	PPV ^c (%)	CUI ^d		
Age 15–64 years																					
Denmark	B34,B97	92.4	89.3	0.83	E	J09, J10	82.4	98.9	0.82	E	J12, J21	23.1	55.4	0.13	P	J09, B34, B97, J10, J12	80.5	95.1	0.77	G	
Iceland	U07.1	97.4	86.5	0.84	E	J09, J10, J11	70.2	62.3	0.44	P	NA					NA					
Malta	U07.1	92.6	92.8	0.86	E	J10, J11	58.1	94.4	0.55	S	J21	10.3	50.0	0.05	P	U07.1, J10, J11	71.6	85.0	0.61	S	
Norway	U07.1	91.1	98.5	0.90	E	J09, J10, J11	92.5	96.7	0.90	E	J12, J20, J21	68.2	18.5	0.13	P	U07.1, J10	77.3	98.1	0.76	G	
Spain	U07.1	96.9	93.1	0.90	E	J09, J10, J11	95.0	99.0	0.94	E	J98	50.0	29.4	0.15	P	U07.1, J10, J12	85.4	88.6	0.76	G	
Age ≥ 65 years																					
Denmark	B34,B97	89.6	98.2	0.88	E	J09, J10	86.4	99.0	0.86	E	J12, J20, J21	27.1	58.9	0.16	P	B34, B97, J09, J10, J12	78.3	96.6	0.76	G	
Iceland	U07.1	96.3	77.7	0.75	G	J09, J10, J11	74.0	59.7	0.44	P	NA					NA					
Malta	U07.1	89.7	88.3	0.79	G	J10, J11	54.8	93.6	0.51	S	J14, J20, J21, J22, J40, J44, J45, J60	73.5	4.1	0.03	P	U07.1, J10, J11	66.3	83.0	0.55	S	
Norway	U07.1	87.4	98.9	0.86	E	J09, J10, J11	91.5	97.7	0.89	E	J12, J20, J21	74.1	22.0	0.16	P	U07.1, J10, J12	82.0	90.7	0.74	G	
Spain	U07.1	95.6	95.8	0.92	E	J09, J10	94.9	98.2	0.93	E	J98	46.9	15.5	0.07	P	U07.1, J10, J12	83.0	93.5	0.78	G	

CUI: clinical utility index; FN: false negative; FP: false positive; ICD-10: 10th edition of the International Classification of Disease codes; PPV: positive predictive value; RSV: respiratory syncytial virus; SARI: severe acute respiratory infection; SARS-CoV-2: severe acute respiratory syndrome coronavirus 2; Se: sensitivity; TP: true positive.

^a ICD-10 codes; U07.1: COVID-19, virus identified. B34.2A: COVID-19, not elsewhere classified (Denmark-specific code). B97.2A: COVID-19 severe acute respiratory syndrome (Denmark-specific code). J03: Acute tonsillitis. J04: Acute laryngitis and tracheitis. J06: Acute upper respiratory infections of multiple and unspecified sites. J09: Influenza due to identified zoonotic or pandemic influenza virus. J10: Influenza due to identified seasonal influenza virus. J11: Influenza due to unidentified influenza viruses. J12: Viral pneumonia, not elsewhere classified. J14: Pneumonia due to *Haemophilus influenzae*. J15: Bacterial pneumonia, not elsewhere classified. J16: Chlamydial pneumonia. J18: Pneumonia, unspecified organism. J20: Acute bronchitis. J21: Acute bronchiolitis. J22: Unspecified acute lower respiratory infection. J40: Bronchitis, not specified as acute or chronic. J44: Other chronic obstructive pulmonary disease. J45: Asthma. J60: Coalworker's pneumoconiosis. Asthma. J91: Pleural effusion in conditions classified elsewhere. J95: Postprocedural respiratory disorders, not elsewhere classified. J96: Respiratory failure, not elsewhere classified. J98: Other respiratory disorders. J99: Respiratory disorders in diseases classified elsewhere.

^b Se (%) = TP / (TP + FN) × 100.

^c PPV (%) = TP / (TP + FP) × 100.

^d CUI = Se × PPV; results are provided both quantitatively and qualitatively, using the following categories: E: excellent utility, G: good utility, S: satisfactory utility, P: poor utility.

across all countries and strata. For the all-ages group, the key codes were J09 and J10 in Spain, J10 and J11 in Malta, and J09, J10 and J11 in Denmark, Norway and Iceland. This suggests that admissions with at least one of these codes correspond to SARI-influenza. In three countries (Denmark, Norway and Spain), the utility achieved across strata was generally excellent (Se ≥ 85%, PPV ≥ 97%, CUI ≥ 0.83 in the all-ages group), with some exceptions among patients younger than 15 years, where the utility was good. Lower accuracy was observed in Iceland (poor utility across all strata) and Malta (poor utility for the 0–4-year-old group and satisfactory utility for the rest of strata).

SARI associated with RSV

For SARI-RSV, the most accurate case definitions across countries included J12 ('Viral pneumonia, not elsewhere classified'), J20 ('Acute bronchitis'), J21 ('Acute bronchiolitis'), and J98 ('Other respiratory disorders'). For the all-ages group, the key codes were J21 in Malta, Norway and Spain, and J12, J20 and J21 in Denmark. Other relevant case definitions for specific strata included J96 (0–4 age group in Denmark),

J98 ('Other respiratory disorders') for the 15–64 and the ≥ 65 age groups in Spain, and J14 ('Pneumonia due to *Haemophilus influenzae*'), J22 ('Unspecified acute lower respiratory infection') and J45 ('Asthma') for the ≥ 65 age group in Malta. However, the utility of all code combinations across countries and strata was generally poor (Se: 35–59%, PPV: 49–71%, CUI: 0.17–0.34 in the all-ages group), except for codes J12 and J21 among children aged 0–4 years in Norway, which achieved satisfactory utility.

SARI associated with any of the three pathogens

For SARI associated with SARS-CoV-2, influenza virus or RSV, the most accurate case definition included U07.1, B34.2A, B97.2A (previously identified for SARS-CoV-2), J09, J10, J11 (identified for influenza virus), and J12, J20, J21 (for RSV). The specific combinations of these codes varied by country. Overall, the utility ranged from satisfactory (Se ≥ 70%, PPV ≥ 78%, CUI ≥ 0.54 for the all-ages group in Malta and Spain) to good (Se ≥ 80%, PPV ≥ 85%, CUI ≥ 0.67 for the all-ages group in Denmark and Norway). However, utility was poor for the 0–4-year-old group in Malta and Spain.

Discussion

Our accuracy assessment identified similar case definitions across countries, with overall satisfactory to excellent utility for SARI associated with SARS-CoV-2 and influenza virus, while accuracy was lower for SARI associated with RSV. For SARI-SARS-CoV-2, a case definition including the four-character COVID-19-specific U07.1 code (B34.2A and B97.2A in Denmark) assigned at admission and discharge provided good to excellent utility across countries. In Canada, studies demonstrated that U07.1 code at discharge indicated excellent utility when using PCR laboratory results as the reference [19,20]. For SARI-influenza, case definitions combining three-character discharge codes J09, J10 and J11 achieved satisfactory to excellent utility, with J10 common to all countries. However, unlike the other countries, utility was poor in Iceland, which may be due to the inclusion of codes at admission as well as at discharge. A similar combination of discharge codes provided results corresponding to good utility in another Canadian study [21]. For SARI-RSV, the key three-character codes were J12, J20 and J21, with J21 common to all countries. However, the overall utility was poor. In Canada, a combination of four-character RSV-specific respiratory and non-respiratory disease codes provided results corresponding to good utility [21]. Furthermore, a Spanish study found that four-character respiratory disease code combinations, including RSV-specific codes, achieved satisfactory utility [22]. In our study, using three-character codes (i.e. non-RSV specific) probably increased Se and mostly decreased PPV, thereby reducing the CUI. Exploring four-character codes for SARI-RSV is therefore recommended. Finally, for SARI associated with any of the three pathogens, a combination of the codes identified for each one achieved satisfactory to good utility, with countries sharing U07.1 (B34.2A+B97.2A in Denmark), J10 and J21.

Importantly, codes with poor utility may still hold value since the qualitative CUI interpretation has not been verified for surveillance and several CUI classifications exist [16,18]. One of them considers CUI values in the range 0.41–0.60 to be of moderate utility, meaning that Iceland's results for SARI-influenza would correspond to moderate rather than poor utility. Furthermore, Davis et al. [23] found that the accuracy of the WHO's 2014 SARI case definition [2] against PCR-confirmed influenza virus and RSV among hospitalised patients yielded accuracy results that corresponded to poor utility [2,23]. To understand the suitability of the identified codes for trend monitoring, we recommend assessing whether weekly rates of pathogen-specific SARI can identify epidemic trends and periods, as defined by weekly test positivity from an external data source.

Cross-country accuracy variations could be attributed to differences in surveillance systems, which we have reported here. These include differences in definitions of hospital admission and readmission, the use of ICD-10 codes assigned at admission or discharge, record

selection processes, time windows for laboratory confirmation, and testing practices. Collectively, these factors contribute to the heterogeneity of national surveillance systems, impacting study participants across countries. This encompasses the characteristics of the reference standard groups, defined by specific ICD-10 codes and available laboratory test results. For instance, since patients without laboratory testing were excluded from the accuracy assessment, settings testing only the most severe cases may exhibit higher test positivity, leading to higher PPV and CUI. Moreover, the availability of laboratory results for influenza virus has been shown to improve the accuracy of ICD-10 coding [24]. Since admission codes are less likely to be informed by laboratory results, they may yield lower accuracy than discharge codes, potentially explaining Iceland's notably lower accuracy for SARI-influenza. The overall effect of these factors on achieved accuracy is unknown, adding complexity to cross-country data interpretation. However, similar codes maximised the Se and PPV across countries.

A limitation of our analysis was that data were retrospectively extracted. As a result, the retained codes primarily represent consolidated discharge diagnoses, which may differ from admission diagnosis codes that could be available and used for real-time trend monitoring [25,26]. Conducting this analysis using weekly data extractions or timestamped information on when codes were assigned would help assess the accuracy of ICD-10 code-based case definitions for real-time surveillance. The lack of consistently structured symptom data in EHR databases impeded their use for defining a SARI reference standard. Instead, we used hospitalisation episodes with respiratory disease or COVID-19 ICD-10 codes and laboratory results for SARS-CoV-2, influenza virus or RSV. True positive SARI pathogen-specific cases were those with laboratory confirmation. This definition was deemed the closest approximation to SARI associated with the pathogens under assessment as it integrates laboratory and clinical information. However, patients not experiencing clinically defined SARI may have been misclassified as true positive pathogen-specific SARI cases (e.g. patients with SARS-CoV-2 confirmation and an unrelated respiratory disease). In addition, we may have missed SARI patients who did not receive respiratory disease or COVID-19 codes but instead were given other codes, such as those for related complications (e.g. hospitalisations with a cardiovascular diagnosis code associated with influenza infection) [11]. Nevertheless, we assume that including non-respiratory disease codes would be unlikely to improve our results due to their expected low PPV. We employed a simplified process to build case definitions, which may be improved by other approaches such as combining more than 10 codes per pathogen, although these may be more burdensome. A Spanish study examined all possible code combinations of a more restricted set of ICD-10 codes for COVID-19 and respiratory diseases presenting seasonal variability, which yielded results similar to ours,

notably for SARI-SARS-CoV-2 and SARI-influenza [22]. Lastly, our study period covered much of the COVID-19 pandemic, which modified the epidemiology of SARI, possibly hospital coding and testing practices, and may have influenced our results [1,11].

Conclusions

Our study applied a methodology for constructing and assessing case definitions based on ICD-10 codes that maximises the Se and PPV for identifying SARI associated with specific pathogens. We found similar diagnosis code sets for SARI associated with SARS-CoV-2, influenza virus, RSV and any of the three pathogens across five countries, despite variations in accuracy level. We recommend replicating this assessment in countries implementing EHR-based SARI surveillance using ICD-10 codes, preferably during non-pandemic periods related to SARI.

ESURE SARI group:

Pello Latasa Zamalloa, Fernando Gonzalez Carril, Liher Imaz Goienetxea, Mikel Ogueta Lana, Eduardo Inchausti Artesero, Aitziber Etxagibel Galdo, Jon Eziolaza Sainz Trapaga, Alba Moya Garcés, Carlota Ruiz de Porras, Luca Basile, Jacobo Mendioroz, María-Teresa Otero-Barrós, Olaia Pérez-Martínez, Ana Sofia Lameiras Azevedo, Miriam López Torrijos, Anthony Nardone

Data availability

See Supplementary Table S4. Additional data may be provided upon request.

Authors' contributions

All authors reviewed the manuscript at different stages of its development. All authors read and approved the final manuscript.

Miguel Angel Sanchez Ruiz: Conceptualisation, Methodology, Formal Analysis, Visualisation, Writing – Original Draft, Writing – Review & Editing.

Diogo F.P. Marques: Conceptualisation, Methodology, Writing – Review & Editing.

Frederikke Kristensen Lomholt: Methodology, Data Curation, Formal Analysis, Writing – Review & Editing

Lasse Skaftve Vestergaard: Methodology, Writing – Review & Editing.

Susana Monge Corella: Methodology, Writing – Review & Editing.

Marcos Lozano Álvarez: Methodology, Data Curation, Formal Analysis, Writing – Review & Editing.

Gudrun Aspelund: Methodology, Writing – Review & Editing.

Marianna Thordardottir: Methodology, Data Curation, Writing – Review & Editing.

Ausra Dziugyte: Methodology, Data Curation, Formal Analysis, Writing – Review & Editing.

John-Paul Cauchi: Methodology, Writing – Review & Editing.

Tjarda M. Boere: Methodology, Data Curation, Formal Analysis, Writing – Review & Editing.

Irene K. Veldhuijzen: Methodology, Writing – Review & Editing.

Elina Seppälä: Methodology, Data Curation, Formal Analysis, Writing – Review & Editing.

Håkon Bøås: Methodology, Formal Analysis, Writing – Review & Editing.

Trine Hessevik Paulsen: Methodology, Writing – Review & Editing.

Ausenda Machado: Methodology, Writing – Review & Editing.

Ana Paula Rodrigues: Methodology, Writing – Review & Editing.

Mariette Hooiveld: Methodology, Writing – Review & Editing.

Luis Alves de Sousa: Project Administration, Methodology, Writing – Review & Editing.

Ana Torres: Project Administration, Methodology, Writing – Review & Editing.

Carlos Carvalho: Project Administration, Methodology, Writing – Review & Editing.

Baltazar Nunes: Supervision, Conceptualisation, Methodology, Formal Analysis, Writing – Original Draft, Writing – Review & Editing.

Pello Latasa Zamalloa: Investigation, Data Curation, Writing – Review & Editing.

Fernando Gonzalez Carril: Investigation, Data Curation, Writing – Review & Editing.

Liher Imaz Goienetxea: Investigation, Data Curation, Writing – Review & Editing.

Mikel Ogueta Lana: Investigation, Data Curation, Writing – Review & Editing.

Eduardo Inchausti Artesero: Investigation, Data Curation, Writing – Review & Editing.

Aitziber Etxagibel Galdo: Investigation, Data Curation, Writing – Review & Editing.

Jon Eziolaza Sainz Trapaga: Investigation, Data Curation, Writing – Review & Editing.

Alba Moya Garcés: Investigation, Data Curation, Writing – Review & Editing.

Carlota Ruiz de Porras: Investigation, Data Curation, Writing – Review & Editing.

Luca Basile: Investigation, Data Curation, Writing – Review & Editing.

Jacobo Mendioroz: Investigation, Data Curation, Writing – Review & Editing.

María Teresa Otero Barrós: Investigation, Data Curation, Writing – Review & Editing.

Olaia Pérez Martínez: Investigation, Data Curation, Writing – Review & Editing.

Ana Sofia Lameiras Azevedo: Investigation, Data Curation, Writing – Review & Editing.

Miriam López Torrijos: Investigation, Data Curation, Writing – Review & Editing.

Anthony Nardone: Funding Acquisition, Conceptualisation, Methodology, Writing – Review & Editing.

Conflict of interest

None declared.

Funding statement

This work was conducted as part of the project “Design and implementation of multinational surveillance systems using routinely collected electronic health records in EU/EEA”, funded by the European Centre for Disease Prevention and Control through a service contract with Epiconcept (specific contract No 3 ECD.14898 ID26495 - implementing framework contract ECDC/2022/003).

Ethical statement

This study was conducted in line with the Declaration of Helsinki. Teams from each participating country are responsible for the national surveillance of SARI. The current analysis was conducted as part of routine surveillance and did not require approval from an ethical review board.

In Denmark, the study was performed as a surveillance study under the authority task of Statens Serum Institut with administrative register data. According to Danish law, ethical approval is not needed for this type of research. In Iceland, the study was performed as a surveillance study under the authority of the Chief Epidemiologist using anonymised administrative register data. According to Icelandic law, ethical approval is not needed for this type of research. In Malta, an ethical review board was not considered necessary because all data analysed were anonymised public health data used for SARI surveillance purposes and were in no way traceable to the patient. In Norway, the legal mandate to perform surveillance of COVID-19 and respiratory infectious diseases causing similar symptoms, without informed consent or need for ethical approval, were governed by the Act on Health and Social Preparedness §2-4. All analyses in the Norwegian SARI surveillance system were performed using the framework of the Emergency preparedness register for COVID-19 (a temporary register established for the monitoring of the COVID-19 pandemic in Norway) and must comply with the data protection impact assessment (DPIA) of this register [10]. In Spain, data extracted by the four participating regions were sent in anonymised form to the centralised analysis team.

Use of artificial intelligence tools

The AI tools Grammarly and ChatGPT were utilised for writing assistance.

References

1. European Centre for Disease Prevention and Control (ECDC). Acute respiratory infections in the EU/EEA: epidemiological update and current public health recommendations.

2. World Health Organization (WHO). Global Influenza Programme. WHO surveillance case definitions for ILI and SARI. Geneva: WHO. [Accessed: 16 Sep 2024]. Available from: <https://www.who.int/teams/global-influenza-programme/surveillance-and-monitoring/case-definitions-for-ili-and-sari>
3. European Centre for Disease Prevention and Control (ECDC). Vaccine Effectiveness, Burden and Impact Studies (VEBIS). Stockholm: ECDC. [Accessed: 10 Jun 2024]. Available from: <https://www.ecdc.europa.eu/en/infectious-disease-topics/related-public-health-topics/immunisation-and-vaccines/vaccine-o>
4. European Centre for Disease Prevention and Control (ECDC). ECDC accession support to the Western Balkans and Türkiye. Stockholm: ECDC; [Accessed: 10 Jun 2024]. Available from: <https://www.ecdc.europa.eu/en/about-ecdc/what-we-do/partners-and-networks/international-cooperation/cooperation-western-balkans-turkiye>
5. European Commission (EC). Design and implementation of multinational surveillance systems using routinely collected electronic health records in EU/EEA. Brussels: EC; 2021. Available from: <https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/tender-details/9819>
6. European Respiratory Virus Surveillance Summary (ERVISS). Overview of respiratory virus epidemiology in the EU/EEA. Stockholm: European Centre for Disease Prevention and Control, Copenhagen: World Health Organization European Region. [Accessed: 16 Sep 2024]. Available from: <https://erviss.org>
7. Marbus SD, van der Hoek W, van Dissel JT, van Gageldonk-Lafeber AB. Experience of establishing severe acute respiratory surveillance in the Netherlands: Evaluation and challenges. *Public Health Pract (Oxf)*. 2020;1:100014. <https://doi.org/10.1016/j.puhip.2020.100014> PMID: 34171043
8. Cauchi JP, Borg ML, Džiugytė A, Attard J, Melillo T, Zahra G, et al. Digitalizing and upgrading severe acute respiratory infections surveillance in Malta: system development. *JMIR Public Health Surveill*. 2022;8(12):e37669. <https://doi.org/10.2196/37669> PMID: 36227157
9. Whittaker R, Toikkanen S, Dean K, Lyngstad TM, Buanes EA, Kløvstad H, et al. A comparison of two registry-based systems for the surveillance of persons hospitalised with COVID-19 in Norway, February 2020 to May 2022. *Euro Surveill*. 2023;28(33):2200888. <https://doi.org/10.2807/1560-7917.ES.2023.28.33.2200888> PMID: 37589591
10. Seppälä E, Bøås H, Dahl J, Stålcrantz J, Stecher M, Tønnessen R, et al. Registry-based surveillance of severe acute respiratory infections in Norway during 2021-2024. *Influenza Other Respir Viruses*. 2025;19(2):e70080. <https://doi.org/10.1111/irv.70080> PMID: 39950571
11. Torres AR, Gómez V, Kislaya I, Rodrigues AP, Fernandes Tavares M, Pereira AC, et al. Monitoring COVID-19 and influenza: the added value of a severe acute respiratory infection surveillance system in Portugal. *Can J Infect Dis Med Microbiol*. 2023;2023:6590011. <https://doi.org/10.1155/2023/6590011> PMID: 36846348
12. Buda S, Tolksdorf K, Schuler E, Kuhlen R, Haas W. Establishing an ICD-10 code based SARI-surveillance in Germany - description of the system and first results from five recent influenza seasons. *BMC Public Health*. 2017;17(1):612. <https://doi.org/10.1186/s12889-017-4515-1> PMID: 28666433
13. El-Khatib Z, Richter L, Reich A, Benka B, Assadian O. Implementation of a surveillance system for severe acute respiratory infections at a tertiary care hospital in Austria: protocol for a retrospective longitudinal feasibility study. *JMIR Res Protoc*. 2023;12(1):e47547. <https://doi.org/10.2196/47547> PMID: 37535414
14. Brady M, Duffy R, Domegan L, Salmon A, Maharjan B, O’Broin C, et al. Establishing severe acute respiratory infection (SARI) surveillance in a sentinel hospital, Ireland, 2021 to 2022. *Euro Surveill*. 2023;28(23):2200740. <https://doi.org/10.2807/1560-7917.ES.2023.28.23.2200740> PMID: 37289427
15. World Health Organization (WHO). ICD-10. International statistical classification of diseases and related health problems 10th revision. Geneva: WHO. [Accessed: 3 May 2024]. Available from: <https://icd.who.int/browse10/2019/en>
16. Mitchell AJ. Sensitivity × PPV is a recognized test called the clinical utility index (CUI+). *Eur J Epidemiol*. 2011;26(3):251-2, author reply 252. <https://doi.org/10.1007/s10654-011-9561-x> PMID: 21442261
17. Østergaard SD, Dinesen PT, Foldager L. Quantifying the value of markers in screening programmes. *Eur J Epidemiol*.

- 2010;25(3):151-4. <https://doi.org/10.1007/s10654-010-9430-z>
PMID: 20151178
18. Mitchell JA. How to: analyse a screening or diagnostic study. *Psycho-oncology.info*; 2009. Available from: https://www.psychononcology.info/PG_analyse_ajmitchell.pdf
 19. Moura CS, Neville A, Liao F, Wen B, Razak F, Roberts S, et al. Validity of hospital diagnostic codes to identify SARS-CoV-2 infections in reference to polymerase chain reaction results: a descriptive study. *CMAJ Open*. 2023;11(5):E982-7. <https://doi.org/10.9778/cmajo.20230033> PMID: 37875313
 20. Moura CS, Morrison LJ, Hohl CM, Grant L, Pilote L, Neville A, et al. Administrative data ICD-10 diagnostic codes identifies most lab-confirmed SARS-CoV-2 admissions but misses many discharged from the Emergency Department. *Sci Rep*. 2024;14(1):6008. <https://doi.org/10.1038/s41598-023-49501-7> PMID: 38472258
 21. Hamilton MA, Calzavara A, Emerson SD, Djebli M, Sundaram ME, Chan AK, et al. Validating International Classification of Disease 10th Revision algorithms for identifying influenza and respiratory syncytial virus hospitalizations. *PLoS One*. 2021;16(1):e0244746. <https://doi.org/10.1371/journal.pone.0244746> PMID: 33411792
 22. Lozano Álvarez M, Moya Garcés A, Imaz Goienetxea L, Lameiras Azevedo AS, Otero Barros MT, Ruiz de Porras Rubio C, et al. Validez de la definición de caso de Infección Respiratoria Aguda Grave basada en códigos CIE-10 en cuatro comunidades autónomas (España, 2021-2023). [Validity of case definitions for Severe Acute Respiratory Infection based on ICD-10 codes in four autonomous communities (Spain, 2021-2023)]. *Rev Esp Salud Publica*. 2025;99:e202505025. Spanish. PMID: 40503697
 23. Davis W, Duque J, Huang QS, Olson N, Grant CC, Newbern EC, et al. Sensitivity and specificity of surveillance case definitions in detection of influenza and respiratory syncytial virus among hospitalized patients, New Zealand, 2012-2016. *J Infect*. 2021;84(2):216-26. <https://doi.org/10.1016/j.jinf.2021.12.012>. PMID: 34953903
 24. Wabe N, Li L, Lindeman R, Post JJ, Dahm MR, Li J, et al. Evaluation of the accuracy of diagnostic coding for influenza compared to laboratory results: the availability of test results before hospital discharge facilitates improved coding accuracy. *BMC Med Inform Decis Mak*. 2021;21(1):168. <https://doi.org/10.1186/s12911-021-01531-9> PMID: 34022851
 25. Mihailovic N, Vasiljevic D, Milicic V, Luketina Sunjka M, Radovanovi S, Milicic B, et al. Discrepancy between admission and discharge diagnoses in central Serbia: analysis by the groups of International Classification of Diseases, 10th revision. *Iran J Public Health*. 2020;49(12):2348-55. <https://doi.org/10.18502/ijph.v49i12.4818> PMID: 34178741
 26. Zikos D, Shrestha A, Fegaras L. Estimation of the mismatch between admission and discharge diagnosis for respiratory patients, and implications on the length of stay and hospital charges. *AMIA Jt Summits Transl Sci Proc*. 2019;2019:192-201. PMID: 31258971

[License, supplementary material and copyright](#)

This is an open-access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0) Licence. You may share and adapt the material, but must give appropriate credit to the source, provide a link to the licence and indicate if changes were made.

Any supplementary material referenced in the article can be found in the online version.

This article is copyright of the authors or their affiliated institutions, 2025.