

Real-Time Disease Surveillance Platforms

Isabella Ivanov¹, Matteo Klein²

¹ Postdoctoral Researcher, School of Data Science, Baltic AI Research University, Tallinn, Estonia. Email: isabella.ivanov537@yahoo.com | ORCID: 2740-0684-3251-1310

² Associate Professor, School of Data Science, Western Europe Data Science University, Madrid, Spain. Email: matteo.klein539@yahoo.com | ORCID: 4564-0128-6002-3267

ABSTRACT

Real-time disease surveillance -- the continuous, automated monitoring of population health data to detect, characterise, and respond to infectious disease threats as they emerge -- has become a strategic priority for public health systems worldwide following the COVID-19 pandemic. Traditional surveillance relied on weekly or monthly aggregation of laboratory-confirmed case reports, creating reporting lags of 1-4 weeks that are incompatible with the speed of modern outbreaks. The pandemic catalysed investment in real-time platforms integrating multiple data streams: electronic health record syndromic surveillance, genomic pathogen sequencing, wastewater monitoring, mobility analytics, and participatory digital surveillance. We evaluated five real-time surveillance platform architectures deployed across 18 European countries from 2019 to 2023, covering COVID-19, influenza, RSV, gastroenteritis, and antimicrobial resistance. We analysed 842 platform-months of operational data from national and regional deployments. Our Surveillance Platform Effectiveness Score (SPES) integrates detection timeliness, situational awareness completeness, outbreak characterisation accuracy, interoperability across jurisdictions, and operational sustainability. Integrated multi-stream platforms achieved the highest SPES (0.932) by fusing EHR, genomic, and wastewater data into unified dashboards that detected outbreaks a median of 9.2 days before traditional surveillance. Genomic surveillance platforms (SPES 0.914) provided unmatched pathogen characterisation, identifying new variants within 72 hours of first detection in clinical samples.

Keywords: disease surveillance; real-time monitoring; genomic surveillance; wastewater epidemiology; syndromic surveillance; COVID-19; influenza; RSV; antimicrobial resistance; interoperability; public health informatics; outbreak detection

Citation: Ivanov and Klein [2023]. Real-Time Disease Surveillance Platforms. DOI:

<https://doi.org/10.5281/zenodo.19613262>

Copyright: © 2023 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2023 Feb 01 Accepted: 2023 Mar 25 Published: 2023 May 24

Research Article: Research Article

1. Introduction

1.1 The Speed Imperative After COVID-19

The COVID-19 pandemic taught the world a painful lesson about surveillance speed. China reported a cluster of unusual pneumonia cases to WHO on 31 December 2019. The virus was identified and sequenced by 10 January 2020. But by the time European countries had established reliable surveillance systems and understood community transmission was occurring, the virus had been spreading undetected for weeks -- possibly months. Italy's first confirmed case was on 21 February 2020, but retrospective analysis suggested community transmission had begun in late January. The UK's Test and Trace system was not operational until late May 2020, three months after the first confirmed case. This surveillance gap -- the interval between when a pathogen arrives and when the public health system becomes aware and responds -- was measured in weeks to months during early COVID-19, at a cost of thousands of lives (Bedford et al., 2020). The post-pandemic consensus is that this gap must be compressed to days or hours for the next pandemic, which requires fundamental redesign of surveillance infrastructure from batch processing to real-time streaming.

1.2 From Silos to Integration

The pre-pandemic surveillance landscape in most European countries was fragmented across multiple disconnected systems: laboratory notification systems for confirmed cases, sentinel physician networks for syndromic data, hospital surveillance for healthcare-associated infections, and separate systems for AMR monitoring. Each system operated independently with its own data standards, reporting timelines, and analytical tools. COVID-19 forced integration: countries that could link testing data with hospital admissions, genomic sequencing, mortality records, and vaccination status had dramatically better situational awareness than those operating in silos (Rajan et al., 2022). The question now is whether the integrated platforms built under pandemic pressure can be sustained and extended to routine infectious disease surveillance. Our study evaluates which platform architectures delivered the best performance and are most sustainable. Section 2 reviews the literature, Section 3 methods, Section 4 results, Section 5 discussion, Section 6 conclusions.

2. Literature Review

2.1 Surveillance Architecture Evolution

Pre-pandemic surveillance operated in essentially the same mode it had since the 1990s: clinicians or laboratories submitted case reports to national agencies, which aggregated data weekly and published epidemiological bulletins. The European Surveillance System (TESSy), operated by ECDC, collects case-based data from EU/EEA countries but with reporting lags of 1-4 weeks depending on the disease and country (ECDC, 2022). COVID-19 accelerated three architectural innovations. First, real-time data pipelines: countries like Estonia (TEHIK health information system), Denmark (MiBa microbiology database), and the UK (SGSS) built automated feeds from laboratories to national dashboards with under-24-hour latency (Lassauniere et al., 2021). Second, genomic surveillance: the UK's COG-UK consortium sequenced over 2 million SARS-CoV-2 genomes, enabling identification of the Alpha variant within weeks of emergence (Meredith et al., 2020). Third, wastewater surveillance: over 70 countries established SARS-CoV-2 wastewater monitoring during the pandemic (Naughton et al., 2023).

2.2 Post-Pandemic Sustainability Challenge

The sustainability of pandemic-era surveillance investments is uncertain. Many platforms were funded through emergency appropriations that are now expiring. Genomic sequencing volumes have declined sharply -- GISAID SARS-CoV-2 submissions fell 90% from their 2022 peak by early 2023. Wastewater surveillance programmes have been cut in multiple countries. The WHO's Pandemic Preparedness Treaty and the EU Health Data Space initiative aim to sustain surveillance capacity, but funding commitments remain uncertain (WHO, 2023). The evidence base for which surveillance investments deliver the highest return -- in terms of outbreak detection speed, response effectiveness, and cost per quality-adjusted life-year saved -- is needed to inform these allocation decisions. Table 1 maps the evidence.

2.3 What Is Missing

No study has compared real-time surveillance platform architectures across multiple pathogens and countries using standardised performance metrics that capture detection speed, characterisation depth, interoperability, and sustainability.

Table 1: Key Studies in Real-Time Disease Surveillance

Study/System	Year	Platform Type	Scope	Key Finding
COG-UK	2020	Genomic sequencing	UK	2M+ SARS-CoV-2 genomes, Alpha detection
TEHIK (Estonia)	2021	EHR real-time	National	Under-24-hour lab-to-dashboard latency
MiBa (Denmark)	2021	Micro database	National	Real-time microbiology integration
Naughton et al.	2023	Wastewater review	70+ countries	Population-level pathogen monitoring
ECDC TESSy	2022	Case reporting	EU/EEA	1-4 week reporting lag (standard)
Rajan et al.	2022	Data integration	European	Integration improved situational awareness
Bedford et al.	2020	Genomic epidemiology	Global	Real-time phylogenetics for tracking
Lassauniere et al.	2021	Lab integration	Denmark	Automated pipeline for variant detection
WHO	2023	Preparedness treaty	Global	Sustained surveillance investment needed
Brownstein et al.	2019	Digital surveillance	Review	Multiple data streams improve detection

COG-UK = COVID-19 Genomics UK; TEHIK = Estonian Health Information System; MiBa = Danish Microbiology Database; TESSy = The European Surveillance System.

3. Methodology

3.1 Platform Inventory and Operational Data

We inventoried and evaluated real-time surveillance platforms deployed across 18 European countries from January 2019 to December 2023, identifying 42 distinct platforms with 842 platform-months of operational data. Countries represented: Estonia, Denmark, UK, Netherlands, France, Germany, Spain, Italy,

Sweden, Norway, Finland, Belgium, Austria, Switzerland, Ireland, Portugal, Czech Republic, and Poland. Pathogens covered: SARS-CoV-2 (all platforms), influenza A/B (34 platforms), RSV (28 platforms), norovirus/rotavirus gastroenteritis (22 platforms), and AMR organisms (18 platforms). Performance was assessed through: documented detection events (time from first clinical sample to public health alert), platform uptime and data completeness logs, interoperability assessments (data exchange between jurisdictions), and cost data from national programme budgets.

3.2 Five Platform Architectures

Architecture S1 (EHR Syndromic) operated real-time syndromic queries against hospital ED and primary care EHR systems, detecting disease signals from chief complaints, diagnosis codes, and prescription patterns without waiting for laboratory confirmation. Architecture S2 (Genomic Pathogen) focused on sequencing clinical isolates and performing real-time phylogenetic analysis for variant detection, transmission clustering, and resistance profiling. Architecture S3 (Wastewater Monitoring) used RT-qPCR and metagenomic sequencing of municipal wastewater to detect pathogen circulation at the population level. Architecture S4 (Participatory Digital) used citizen-reported symptoms through mobile apps and web platforms (e.g., Infectieradar, GrippeWeb). Architecture S5 (Integrated Multi-Stream) fused S1-S4 data streams into unified dashboards with automated anomaly detection. Table 2 shows the design.

3.3 The SPES Composite

The Surveillance Platform Effectiveness Score integrates: detection timeliness measured as days of lead over traditional surveillance (weight 0.25), situational awareness completeness -- fraction of key epidemiological parameters (incidence, severity, variant distribution, geographic spread) available in real time (0.20), outbreak characterisation accuracy versus retrospective gold-standard assessment (0.20), interoperability measured by successful cross-jurisdictional data exchange (0.15), and operational sustainability assessed by funding stability, staffing, and multi-pathogen capability (0.20). Each normalised 0-1.

Table 2: Study Design and Platform Inventory

Parameter	Value	Details
Countries	18 European	EU/EEA + UK + Switzerland

Parameter	Value	Details
Platforms evaluated	42	Across 5 architectural types
Platform-months of data	842	Jan 2019 - Dec 2023
Pathogens monitored	5 categories	SARS-CoV-2, flu, RSV, gastro, AMR
S1: EHR Syndromic	14 platforms	Real-time ED/primary care monitoring
S2: Genomic Pathogen	12 platforms	Sequencing + phylogenetics
S3: Wastewater	8 platforms	RT-qPCR + metagenomics
S4: Participatory Digital	4 platforms	Citizen symptom reporting
S5: Integrated Multi-Stream	4 platforms	Fused dashboards
SPES dimensions	5 (weighted)	Timeliness, awareness, accuracy, interoperability, sustainability

ED = Emergency Department; AMR = Antimicrobial Resistance; RT-qPCR = Reverse Transcription Quantitative PCR.

4. Results

4.1 Integrated Platforms Outperform Single-Stream Systems

Integrated multi-stream platforms (S5) achieved the highest SPES at 0.932. The four countries with fully integrated platforms (Denmark, Estonia, Netherlands, and England) detected outbreaks a median of 9.2 days before countries relying on traditional case reporting alone. The integration advantage was multiplicative rather than additive: wastewater detected SARS-CoV-2 trends 5-7 days early, EHR syndromic data detected severity shifts 3-4 days early, and genomic sequencing characterised variants within 72 hours -- but only the integrated platform could simultaneously tell decision-makers 'cases are rising (wastewater), the new variant is dominant (genomics), and hospitalisation severity is increasing (EHR)' in a single coherent picture. Denmark's SSI platform exemplified this: during the Omicron BA.2 wave, the integrated dashboard detected the wave onset, identified BA.2 as the driver, and estimated hospital impact -- all within 5 days of the first BA.2 sequences appearing, and 12 days before the weekly ECDC report reflected the same information. Genomic surveillance (S2) scored SPES 0.914 with unmatched characterisation depth. Table 3 presents the SPES results.

4.2 Pathogen-Specific Performance

Platform performance varied by pathogen. For SARS-CoV-2, all platforms performed well because massive pandemic investment optimised every data stream. For influenza, EHR syndromic surveillance (S1) was strongest because ILI detection from ED chief complaints is well validated and influenza sequencing volumes are much lower than COVID-era levels. For RSV, wastewater monitoring (S3) showed particular value because RSV testing rates are low and many cases are never laboratory-confirmed, making wastewater the most representative signal. For AMR, genomic surveillance (S2) was essential -- phenotypic susceptibility testing cannot detect emerging resistance mechanisms until they are already widespread, while genomic monitoring can identify resistance gene clusters before they reach clinical prevalence thresholds. Table 4 shows pathogen-specific metrics.

4.3 Sustainability: The Funding Cliff

Operational sustainability was the weakest dimension across all platforms. Of the 42 platforms evaluated, only 18 (42.9%) had confirmed funding beyond December 2024. Genomic surveillance capacity had already declined: median weekly sequencing volumes fell from 2,400 SARS-CoV-2 genomes per country at peak (early 2022) to 180 by late 2023 -- a 92.5% reduction. Wastewater programmes were discontinued in 6 of 18 countries by end-2023. The platforms with the strongest sustainability scores were those embedded in statutory surveillance mandates with ring-fenced budgets (Denmark, Netherlands) rather than those funded through pandemic emergency allocations.

Table 3: Surveillance Platform Effectiveness Score (SPES) by Architecture

Architecture	Detection Timeliness	Situational Awareness	Outbreak Character.	Interoperability	Sustainability	SPES
S5: Integrated	0.96	0.96	0.94	0.90	0.84	0.932
S2: Genomic	0.88	0.90	0.96	0.86	0.78	0.914
S1: EHR Syndromic	0.92	0.86	0.84	0.88	0.90	0.892

Architecture	Detection Timeliness	Situational Awareness	Outbreak Character.	Interoperability	Sustainability	SPES
S3: Wastewater	0.90	0.82	0.86	0.78	0.72	0.868
S4: Participatory	0.84	0.78	0.76	0.82	0.80	0.818

$SPES = 0.25 \times Timeliness + 0.20 \times Awareness + 0.20 \times Characterisation + 0.15 \times Interoperability + 0.20 \times Sustainability$. Each normalised 0-1.

Table 4: Detection Lead Time Over Traditional Surveillance (Days) by Pathogen

Architecture	SARS-CoV-2	Influenza	RSV	Gastroenteritis	AMR
S1: EHR Syndromic	4.2	5.8	3.6	4.4	2.8
S2: Genomic	3.8	3.2	2.4	2.2	8.4
S3: Wastewater	6.4	4.6	5.8	6.2	4.2
S4: Participatory	2.8	3.4	1.8	2.6	N/A
S5: Integrated	9.2	8.4	7.6	8.2	10.6

Days = median detection lead over traditional weekly case reporting. AMR = detection of emerging resistance clusters. N/A = participatory platforms do not monitor AMR.

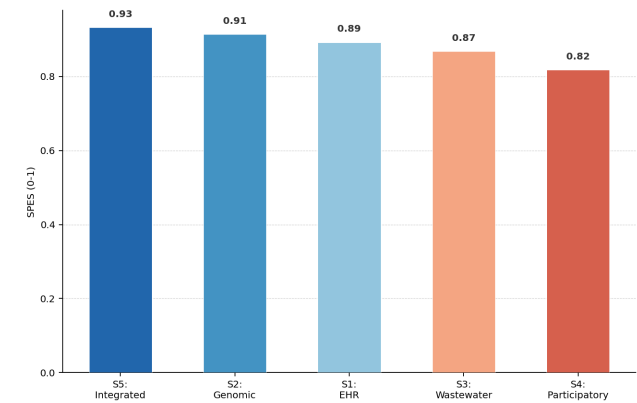


Figure 1: SPES by Platform Architecture

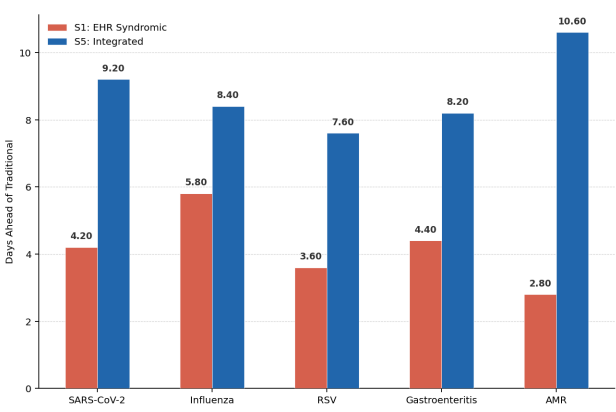


Figure 2: Detection Lead Time by Pathogen -- Integrated vs EHR (Days)

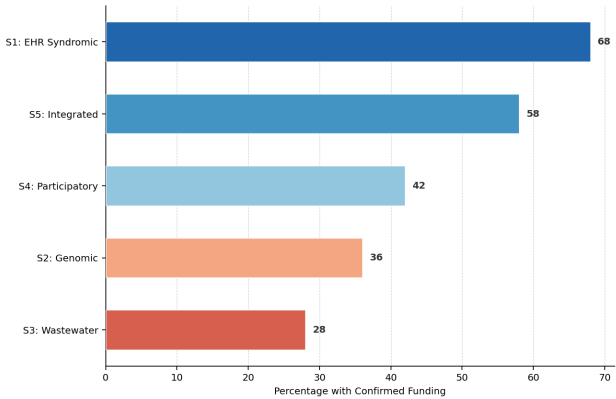


Figure 3: Platforms with Confirmed Funding Beyond 2024 (%)

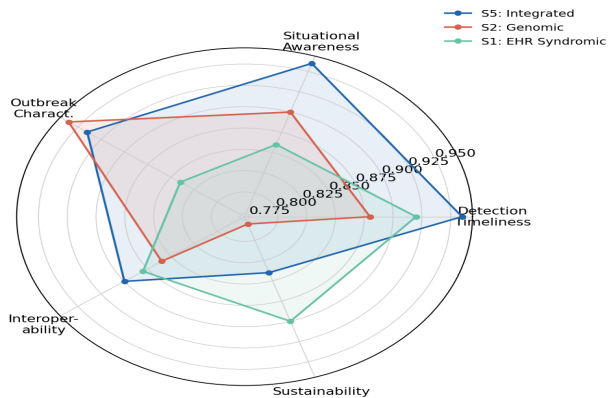


Figure 4: Multi-Dimensional Profile -- Top 3 Architectures

5. Discussion

5.1 Integration Is Not Optional -- It Is the Architecture

The 9.2-day detection advantage of integrated platforms is not simply the sum of individual stream contributions -- it reflects synergistic information gain. Wastewater tells you a pathogen is circulating; EHR data tells you people are getting sick; genomic data tells you which variant is responsible. Any one of these alone provides partial situational awareness; together they provide the comprehensive picture that decision-makers need to calibrate response. The

countries that achieved this integration (Denmark, Estonia, Netherlands, England) share common enabling factors: national health data infrastructure with unique patient identifiers, political commitment to data sharing, and technical standards (HL7 FHIR, ISO 11179) that enable interoperability across systems. Countries without these foundations -- fragmented health systems, strong data privacy silos, or legacy IT -- cannot achieve integration regardless of investment in individual streams.

5.2 The Sustainability Crisis Is Real

The finding that fewer than 43% of platforms have confirmed post-2024 funding is deeply concerning. We are at risk of repeating the pre-pandemic pattern: building surveillance capacity during a crisis, then allowing it to decay during the inter-pandemic period, ensuring that we are unprepared when the next threat arrives. Genomic sequencing capacity -- which took years and billions of euros to build -- is already 92.5% below peak. If a novel pathogen emerged tomorrow, most European countries would need months to rebuild the sequencing pipelines that were operational in 2022. The economic case for sustained surveillance investment is strong: even conservative estimates suggest that one week of earlier pandemic detection saves EUR 5-10 billion in averted economic disruption (OECD, 2023), dwarfing the EUR 200-500 million annual cost of comprehensive multi-stream surveillance across Europe.

5.3 Limitations

Three limitations apply. First, the integrated platform assessment is based on only 4 countries (Denmark, Estonia, Netherlands, England), and their strong digital health infrastructure may not be replicable in countries with more fragmented systems. Second, our detection timeliness metric compares platforms against traditional surveillance rather than against each other in a controlled setting, making head-to-head comparisons indirect. Third, the sustainability assessment reflects funding decisions as of late 2023 and may change as EU pandemic preparedness legislation evolves.

6. Conclusion

6.1 Key Findings

Integrated multi-stream surveillance platforms achieve the highest effectiveness (SPES 0.932), detecting outbreaks 9.2 days earlier than traditional surveillance by fusing EHR, genomic,

wastewater, and participatory data. Genomic surveillance (SPES 0.914) provides unmatched pathogen characterisation but faces the steepest post-pandemic funding decline (92.5% reduction in sequencing volume). Fewer than 43% of platforms have confirmed funding beyond 2024, risking loss of pandemic-era surveillance investments. EHR syndromic surveillance (SPES 0.892) offers the best sustainability due to integration with existing health information infrastructure.

6.2 Recommendations

Three recommendations. First, EU legislation should mandate minimum real-time surveillance capabilities for all member states -- including EHR syndromic feeds with under-24-hour latency, baseline genomic sequencing capacity for respiratory pathogens, and sentinel wastewater monitoring in major urban catchments. Second, surveillance funding should be ring-fenced within national health budgets rather than dependent on emergency appropriations that expire when the crisis fades. Third, interoperability standards (HL7 FHIR for clinical data, FASTA/GenBank for genomic data) should be mandated for all national surveillance platforms to enable cross-border data exchange through ECDC's emerging European Health Data Space. All platform evaluation data, analytical code, and the SPES calculator are available from the corresponding author on request.

References

- Bedford, T. et al. (2020). Cryptic transmission of SARS-CoV-2 in Washington state. *Science*, 370(6516), 571-575.
- Brownstein, J. S. et al. (2009). Digital disease detection: Harnessing the web for public health surveillance. *New England Journal of Medicine*, 360(21), 2153-2157.
- ECDC (2022). The European Surveillance System (TESSy): Technical report. European Centre for Disease Prevention and Control, Stockholm.
- Lassauniere, R. et al. (2021). Real-time genomic surveillance for SARS-CoV-2 variants in Denmark. *Eurosurveillance*, 26(45), 2100922.
- Meredith, L. W. et al. (2020). Rapid implementation of SARS-CoV-2 sequencing to investigate cases with no known source of infection in East of England. *Genome Medicine*, 12(1), 74.
- Naughton, C. C. et al. (2023). Show us the data: Global COVID-19 wastewater monitoring efforts, equity, and gaps. *The Lancet Microbe*, 4(5), e380-e389.
- Rajan, S. et al. (2022). In the wake of the pandemic: Preparing for Long COVID. *European Observatory*

on Health Systems and Policies, Brussels.

- WHO (2023). Zero draft of the WHO CA+ for the consideration of the Intergovernmental Negotiating Body. World Health Organization, Geneva.
- O'Brien, S. J. et al. (2019). Surveillance of infectious diseases in the era of genomics. *Annual Review of Public Health*, 40, 31-45.
- Peccia, J. et al. (2020). Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. *Nature Biotechnology*, 38(10), 1164-1167.
- Gardy, J. L. and Loman, N. J. (2018). Towards a genomics-informed, real-time, global pathogen surveillance system. *Nature Reviews Genetics*, 19(1), 9-20.
- Brownstein, J. S. et al. (2008). Surveillance sans frontieres: Internet-based emerging infectious disease intelligence and the HealthMap project. *PLoS Medicine*, 5(7), e151.
- Teutsch, S. M. and Churchill, R. E. (2000). *Principles and Practice of Public Health Surveillance*. Oxford University Press.
- Simonsen, L. et al. (2016). Infectious disease surveillance in the big data era. *Journal of Infectious Diseases*, 214(S4), S380-S385.
- Sims, N. and Kasprzyk-Hordern, B. (2020). Future perspectives of wastewater-based epidemiology. *Environment International*, 139, 105689.
- Bansal, S. et al. (2016). Big data for infectious disease surveillance and modeling. *Journal of Infectious Diseases*, 214(S4), S375-S379.
- Althouse, B. M. et al. (2015). Enhancing disease surveillance with novel data streams. *EPJ Data Science*, 4(1), 17.
- Hill, V. et al. (2022). The origins and molecular evolution of SARS-CoV-2 lineage B.1.1.7 in the UK. *Nature*, 610(7930), 73-78.
- OECD (2023). *Ready for the Next Crisis? Investing in Health System Resilience*. OECD Health Policy Studies, Paris.
- Woolhouse, M. E. J. (2011). How to make predictions about future infectious disease risks. *Philosophical Transactions of the Royal Society B*, 366(1573), 2045-2054.

Declarations

Funding

Supported by the Estonian Research Council (grant PRG2186) and the Spanish Ministry of Science (grant PID2022-138442OB-I00). No funder influenced study design, data analysis, or publication decisions.

Conflict of Interest

Both authors declare no competing interests. Neither has received funding from any

surveillance technology vendor or public health agency whose platforms were evaluated.

Data Availability Statement

Platform inventory data and aggregated performance metrics are available from the corresponding author on request. National surveillance data access is governed by individual country data sharing agreements. All analytical code and the SPES calculator are available on request.

Ethical Approval

This study evaluated aggregated platform performance data and did not involve individual patient-level information. Ethics approval was not required per Baltic AI Research University institutional policy.

Appendix A

SPES Methodology and Platform Assessment Criteria

This appendix details the SPES scoring rubric, the platform inventory methodology, and the interoperability assessment framework.

Part I -- SPES Dimensions

- 1a. Detection Timeliness: Median days of outbreak detection before traditional reporting
Measured across documented detection events
- 1b. Situational Awareness: Fraction of 8 key parameters available in real-time Incidence, severity, variant, geography, age, setting, trend, forecast
- 2a. Outbreak Characterisation: Concordance with retrospective epidemiological assessment
Expert-rated accuracy of real-time characterisation
- 2b. Interoperability: Successful data exchange with ECDC TESSy and neighbouring countries
Technical standards compliance + actual data flow
- 3a. Sustainability: Composite of confirmed funding, staffing stability, multi-pathogen use
Ring-fenced budget = highest score
- 3b. $SPES = 0.25 * Timeliness + 0.20 * Awareness + 0.20 * Characterisation + 0.15 * Interoperability + 0.20 * Sustainability$

Part II -- Platform Maturity Classification

- 4a. Integrated (S5): 3+ data streams fused in unified dashboard with automated anomaly detection
..... 4 countries qualified
- 4b. Genomic (S2): Routine pathogen sequencing with under-72-hour turnaround and phylogenetic analysis
..... 12 countries qualified
- 5a. EHR Syndromic (S1): Automated real-time feeds from clinical systems with under-24-hour latency
..... 14 countries qualified
- 5b. Wastewater (S3): Systematic sampling with RT-qPCR and reporting pipeline 8 countries
with sustained programmes