Oral Sentinel: Clinical Weak Signals as a Pre-Diagnostic Layer for Pandemic Early Warning

A policy and systems analysis for distributed bio-surveillance

Filipa Martins

With Apart Research

[1] Research conducted at the AIxBio Hackathon, April 2026

Abstract

Early warning systems for biological threats typically rely on laboratory, environmental, or digital surveillance, often detecting signals only after transmission is underway. This project explores whether routine clinical observation, specifically in public oral health services, can function as an earlier, underutilized detection layer. We propose Oral Sentinel, a minimal weak-signal reporting concept that captures atypical oral and mucosal findings as potential early indicators of emerging biological events. Rather than building a full technical system, the project defines a practitioner-facing reporting logic, focusing on signal framing, usability, and integration with existing workflows. The contribution is conceptual and design-oriented: mapping how neglected clinical infrastructures could augment current surveillance architectures. We discuss feasibility, limitations, and the role of qualitative practitioner knowledge in early detection ecosystems. The central reframing is this: improving early warning may depend as much on where signal is captured as on how it is modeled.

Hackathon Contribution

During the hackathon, this project focused on translating a neglected early-warning hypothesis into a pilotable system design. Specifically:

- Framed routine oral clinical observations as a potential pre-diagnostic signal layer
- Developed a minimal reporting logic for weak-signal capture in practice
- Positioned the concept within existing pandemic early warning architectures
- Identified key limitations, risks, and implementation challenges

This work does not present a deployed system, but a structured proposal for integrating practitioner-level signals into biosurveillance architectures.

1. Introduction

Current biosurveillance systems assume signals already exist; this work focuses on enabling earlier signal generation in under-instrumented clinical environments. This matters especially in an AI x biosecurity landscape, where advances in AI may compress the time available to detect unusual, engineered, or deliberately caused biological events. But this leaves a major design question underexplored: what routine infrastructures already encounter weak signal before formal surveillance systems know how to see it? This work reframes early warning from better detection models to earlier signal capture within neglected clinical infrastructures.

This report argues that public oral health services may represent one such neglected surveillance-relevant infrastructure. Dentists and oral hygienists, particularly in public, preventive, and school-linked settings, routinely observe oral and mucosal findings, interact with populations outside acute illness pathways, and operate within distributed care environments largely invisible to epidemic intelligence systems. The paper proposes that this infrastructure could function as a complementary sentinel layer, not by diagnosing outbreaks directly, but by contributing structured weak-signal reporting during routine care and supporting outbreak-linked surveillance workflows under defined public health conditions.

The report develops a two-tier model: (1) low-burden anomaly capture embedded into routine oral health workflow, and (2) targeted activation for decentralized surveillance support when triggered by public health authorities. It further argues that the most useful role for AI in this context may not be only downstream in anomaly detection, but also upstream at the point of capture, helping structure observation, improve documentation, and making weak clinical signal more legible in the first place.

2. Research Question

This report asks whether public oral health services could function as a viable sentinel layer in epidemic early warning systems, and what technical, governance, and workflow conditions would be required for such integration to be useful, proportionate, and operationally realistic.

More specifically, it explores whether a distributed oral health workforce, including public dentists and oral hygienists, could contribute to epidemic early warning through: (1) low-burden reporting of unusual oral or contextual clinical patterns, and (2) outbreak-activated participation in decentralized surveillance workflows, such as saliva-based sample collection where appropriate.

The report investigates a narrower systems question:

Can an underused clinical infrastructure provide useful weak-signal visibility when integrated into a broader public health surveillance architecture?

This is approached as a feasibility-oriented policy and systems analysis, rather than an empirical effectiveness study.

3. Why This Matters Now

3.1. Epidemic surveillance is strongest once disease becomes clinically visible

Contemporary epidemic surveillance systems are largely organized around a familiar set of entry points: emergency departments, primary care, laboratory reporting, and formal notification systems. These channels are indispensable for confirming cases, monitoring disease burden, and coordinating response. However, they are not equally well-suited to early, diffuse, low-specificity signal detection. Many existing surveillance pathways become most informative only once illness has become sufficiently visible to enter mainstream medical attention, through acute symptoms, diagnostic testing, or healthcare-seeking behaviour. Public health systems may be well-equipped to detect disease once it becomes clinically legible, but less equipped to notice weak signals while they are still socially and geographically diffuse.

3.2. Earlier weak-signal visibility becomes more important under rising biological risk

This challenge becomes more salient in a changing biosecurity landscape. As AI expands access to biological reasoning, modeling, sequence analysis, and operational support, concern is growing that some forms of biological risk may become faster-moving, more distributed, or more difficult to anticipate through traditional pathways alone. It is no longer sufficient to ask only which model should detect the outbreak, or which central database should receive the alert. It also becomes necessary to ask: which already-existing human, clinical, and community infrastructures might produce weak but useful signal before escalation?

3.3. The neglected design problem: overlooked infrastructures

A large share of surveillance innovation focuses on new diagnostic technologies, AI-enabled pattern detection, digital epidemiology, and centralized data integration. These are important. But they can obscure a more basic systems design question: what infrastructures are already present, already distributed, and already in contact with the population, but not currently treated as surveillance-relevant? This report argues that public oral health services may be one such overlooked infrastructure, not proposed as a replacement for existing surveillance systems, but as a complementary layer with distinct observational reach.

4. The Overlooked Infrastructure: Why Public Oral Health Services

4.1. Contact outside acute illness pathways

Routine oral health encounters often involve people who are not presenting primarily for systemic illness, making these settings structurally different from acute medical pathways. From a surveillance perspective, this matters because they provide access to different kinds of contact and observation than systems built around overt medical presentation. In public health systems with school-linked oral health programs, oral hygienists working in preventive settings may encounter children and school populations in ways that create distinct public health visibility, particularly in community-facing environments often weakly integrated into formal surveillance systems.

4.2. The oral cavity as a weak-signal observation site

Public oral health services involve direct observation of the oral cavity and surrounding tissues, an anatomical site that may, in some cases, display clinically meaningful changes associated with infectious or systemic processes. Potentially relevant observations may include: unexplained ulcerative lesions, vesicular or blister-like presentations, unusual inflammatory or hemorrhagic mucosal patterns, or abrupt local increases in similar presentations across time. The key claim is not that oral health professionals can identify emerging pathogens clinically, but that routine oral examination may provide access to weak but clinically legible signal that is currently under-structured and underused in surveillance design.

4.3. Operational fit: structured capture and outbreak activation

Unlike many hypothetical early-warning ideas, oral health services already provide a distributed workforce, routine face-to-face encounters, existing documentation practices, and a care environment compatible with low-burden structured signal capture. A viable sentinel model would depend less on creating a new surveillance platform and more on embedding minimal structured capture inside existing workflows, for example through a lightweight module within existing oral health information systems, allowing clinicians to log predefined anomaly categories through checkbox inputs rather than narrative reporting.

5. Hypothesis and Contribution

5.1. Core hypothesis

Public oral health services could function as a useful low-burden sentinel layer in epidemic early warning systems by contributing structured weak-signal reporting during routine care, and by supporting decentralized surveillance workflows when activated under outbreak conditions. The plausibility of this hypothesis does not depend on high diagnostic specificity at the level of individual observations. Its potential value lies in distribution, repetition, structure, and aggregation, namely signal density rather than individual accuracy.

5.2. Supporting hypotheses

H2 — Workflow integration: A narrowly defined set of reportable anomaly categories can be integrated into routine oral health workflows with minimal additional burden, particularly if embedded as structured, low-friction inputs within existing systems.

H3 — Weak-signal aggregation: Low-specificity oral or contextual signals may become more useful when interpreted as aggregated patterns across time and geography, rather than as isolated clinical findings.

H4 — Implementation constraints: The main barriers to integrating oral health services into surveillance systems are likely to be governance, interoperability, and workflow design constraints, rather than lack of technical capability alone.

5.3. Contribution

The contribution of this report is not empirical proof, but a structured feasibility argument: a reframing of public oral health services as overlooked sentinel infrastructure; a plausible two-tier model for how that infrastructure might be used; and an implementation-oriented analysis of why such integration is not already operational. This reframing suggests that improving early warning may depend as much on where signal is captured as on how it is modeled.

6. Proposed System Model

This report proposes a two-tier integration model for incorporating public oral health services into epidemic early warning architecture. The model is intentionally designed to be low-burden, non-diagnostic, proportional to risk, and compatible with routine public-sector workflow.

6.1. Tier 1 — Passive anomaly reporting (routine low-burden sentinel sensing)

The first layer is designed for ordinary time, outside any confirmed outbreak context. Its function is simple: to allow oral health professionals to log a small number of predefined anomaly categories during routine encounters through a structured and low-friction workflow. Possible reportable categories include: unexplained ulcerative lesions, vesicular or blister-like oral presentations, unusual inflammatory or hemorrhagic mucosal findings, and sudden apparent increases in similar presentations over a short time window.

Workflow logic: Signal capture should be embedded into existing oral health information systems through checkbox-based anomaly logging. A structured anomaly layer could be further strengthened by lightweight clinical decision support embedded into the same workflow, prompting the clinician through related observations if a category is triggered. The AI layer here is not surveillance infrastructure; it is clinical support that makes structured observation easier to perform and more consistent across clinicians.

6.2. Tier 2 — Outbreak-activated surveillance support

The second layer is designed for activation under clearly defined public health conditions. Activation would only be justified where three criteria are met: (1) a public health trigger exists (suspected local cluster, confirmed outbreak, or enhanced local surveillance need); (2) the pathogen or surveillance objective is compatible with low-friction oral or saliva-based sampling; and (3) the purpose is to extend decentralized surveillance reach, not to replace standard diagnostic pathways. Tier 2 is not a standing surveillance function; it is a contingency-use extension layer.

6.3. Design principles for a viable pilot

Minimal burden: Any added task must take seconds, not minutes.

Clinical legibility: Reporting categories must make sense to clinicians in the context of actual care.

Local usefulness: The system should ideally improve routine clinical structure, not merely extract data for external use.

Human review: Signal interpretation and escalation should remain under public health oversight.

Proportionality: Narrow in ordinary time; intensification only under explicit outbreak conditions.

Professional role realism: Different professional groups included according to the actual structure of public oral health services.

This model is intentionally designed to be pilotable within existing public oral health systems without requiring new standalone infrastructure, enabling immediate feasibility testing in bounded settings.

7. Why This Is Not Already Operational

The barriers to integrating oral health services into surveillance architecture are primarily institutional, not technical. Key challenges include: (1) interoperability gaps, meaning oral health records are not meaningfully integrated into public health surveillance architecture; (2) absence of minimal reporting standards, as no shared taxonomy exists for what counts as a reportable weak signal; (3) workflow burden, since in public health systems, additional reporting responsibilities only become real if they survive contact with actual service conditions; (4) governance ambiguity, as Tier 2 cannot depend on informal ad hoc clinical judgment; it requires pre-defined activation authority and clear institutional ownership; and (5) legal and ethical proportionality, as any model must be designed around minimal data collection, limited and purpose-specific use, and explicit governance from the start.

The central implementation question is not whether signal exists, but whether institutions can be made capable of noticing, structuring, and sustaining it.

8. Expected Results

The pilot is not designed to demonstrate outbreak detection directly, but to establish whether structured clinical weak signals can be captured, aggregated, and made operationally usable within existing workflows, the necessary preconditions for a system that could scale:

Feasibility: A minimal anomaly capture layer can be embedded into routine oral health workflow with acceptable burden, particularly if the taxonomy remains narrow and reporting is structured.

Signal construction: Routine encounters can generate low-specificity but non-random signal when observations are structured consistently enough to be aggregated. The pilot period itself constitutes the primary mechanism for baseline construction, a meaningful result in its own right.

Implementation findings: The main constraints would likely emerge less from lack of signal than from institutional and workflow friction, including uneven uptake, variability in category use, and dependency on whether the task feels clinically useful or administratively extractive.

Role of AI: AI becomes useful only once the underlying clinical and reporting infrastructure is made structured enough to produce signal in the first place, not before.

Secondary clinical results: More systematic observation of mucosal and soft-tissue findings may improve clinical quality independently, including earlier detection of oral pathology such as oral cancer or systemic disease manifestations.

9. Discussion

9.1. The bottleneck may lie upstream of analytics

A common assumption in AI x biosecurity is that better early warning depends primarily on better models. The oral health case suggests that the deeper bottleneck is often upstream: potentially relevant weak signal is never made structured enough, visible enough, or governable enough to become legible. That shifts the core design question from 'what model should detect outbreak-relevant patterns?' to 'which routine infrastructures already encounter weak signal, and what would it take to make that signal usable?'

9.2. The most useful AI may sit at the point of capture

The most useful role for AI may not be only downstream in anomaly detection or aggregation. It may be at the point of capture, helping clinicians notice, structure, or consistently document weak signal. A lightweight AI-supported layer that improves structured observation at point of care may be more valuable than a more sophisticated model applied later to poorly captured data. This reframes the role of AI from retrospective detector toward clinical and infrastructural enabler.

9.3. Durable surveillance may depend on improving care, not just extracting data

Surveillance systems are unlikely to persist if they are designed only as extractive layers above routine care. What makes this proposal more plausible is that the same workflow might also improve something locally meaningful, including more structured mucosal observation, better follow-up of suspicious lesions, and greater attention to oral manifestations that are often clinically under-recorded. Durable systems survive when they are institutionally livable.

9.4. A broader class of neglected infrastructures

Oral health services are unlikely to be the only routine system with underused surveillance value. They may be one example of a broader class of neglected observational infrastructures, meaning service environments that already encounter weak signal, but are not treated as epistemically relevant because they sit outside traditional surveillance imagination. Some of the most useful innovations in biosurveillance may come not from entirely new sensors, but from recognizing that existing institutions already contain observational capacity that has never been properly structured, integrated, or governed.

10. Conclusion

This report has argued that public oral health services may represent a plausible but currently underused sentinel layer in epidemic early warning systems. Their potential value does not lie in diagnostic certainty, but in a more modest and operationally relevant possibility: that routine oral health encounters may generate enough structured weak signal to become useful when integrated into a broader surveillance architecture.

The central contribution is not only a proposal about oral health services. It is a broader claim about AI x biosecurity: that some of the most useful surveillance innovation may come not from new detectors alone, but from recognizing and structuring signal that already exists inside neglected routine infrastructures. This work reframes early warning from detection models to signal capture infrastructure, a shift with implications beyond oral health. Public oral health services appear sufficiently plausible, underused, and operationally relevant in some contexts to justify bounded pilot testing.

References

1. Hughes HE, et al. Emergency department syndromic surveillance systems: a systematic review. *BMC Public Health.* 2020;20:1891.

2. Torres-Urquidy MH, et al. Detection of disease outbreaks by the use of oral manifestations. *J Dental Research.* 2009;88(1):89–94.

3. Ruiz-Mojica CA, Brizuela M. Viral infections of the oral mucosa. In: *StatPearls.* StatPearls Publishing; 2023.

4. Jensen TW, et al. Saliva has high sensitivity and specificity for detecting SARS-CoV-2. *Diagnostics.* 2025;15(15):1918.

5. Tran L, Tran P. The potential of saliva as a sensitive sample type for detection of respiratory pathogens. *The Lancet Microbe.* 2023;4(9):e739–e748.

6. Direção-Geral da Saúde (DGS). Programa Nacional de Promoção da Saúde Oral (PNPSO). Lisboa: DGS. Accessed April 2026.

7. Johns Hopkins Center for Health Security. AI x Biosecurity: Managing Biological Risks in the Age of Artificial Intelligence. Accessed April 2026.

8. Kalenderian E, et al. COVID-19 provides an opportunity for integration of dentistry into the health informatics system. *JADA.* 2021;152(12):952–958.

9. Torres-Urquidy MH. Biosurveillance and Dentistry. In: Acharya A, et al., eds. *Integration of Medical and Dental Care and Patient Data.* Springer; 2019.

Appendix: Limitations and Dual-Use Considerations

A. Limitations

Low specificity of oral findings. Many potentially relevant oral findings are highly non-specific. Ulcerative, vesicular, or hemorrhagic oral lesions can arise from a wide range of causes unrelated to outbreak dynamics. Their possible value lies only in potential contribution as weak signal in aggregate, not as standalone clinical alerts. A poorly designed model could easily mistake common background noise for surveillance-relevant signal.

Risk of false positives and alert fatigue. Because the model relies on weak and low-specificity observations, it carries an inherent risk of generating false positives, spurious clusters, or non-actionable alerts. This makes conservative threshold design and mandatory human review essential.

Uneven population coverage. Public oral health services do not provide equal access across all populations; any sentinel role would remain partial. However, in systems with school-linked oral health programs, these may provide relatively broad access to pediatric populations, a potentially important surveillance advantage in child-facing contexts.

Workflow variability and observer heterogeneity. The model depends on relatively consistent observation and reporting across professionals and settings. Different clinicians may vary in how they examine mucosal findings, interpret borderline lesions, or engage with structured categories. This heterogeneity directly affects signal quality and should be treated as a central design and evaluation problem.

Risk of over-medicalizing routine care. There is a conceptual and ethical risk in turning routine oral health encounters into surveillance-adjacent events. Any implementation must remain clearly bounded, proportionate, and clinically subordinate to the primary purpose of care.

Risks associated with AI-supported structuring. AI-supported prompts could over-standardize clinician attention or create false confidence in low-certainty observations. At system level, AI-assisted aggregation may create a misleading sense of objectivity if underlying data remain sparse or noisy. The role of AI in this model should remain supportive and bounded, not authoritative.

Saliva-based surveillance has narrower value than may appear. Its utility varies by pathogen, collection quality may be inconsistent, and its operational value lies in ease and access rather than universal diagnostic superiority. Saliva should not be treated as the central justification for integrating oral health services into surveillance architecture.

The baseline validation problem. The usefulness of the system cannot be demonstrated before a baseline exists, and a baseline cannot exist before the system is implemented. This means a pilot would need to be evaluated primarily through process metrics in its early phase.

B. Dual-Use Considerations

This system is explicitly not designed for real-time outbreak detection or automated alerting, but for structured signal generation under public health oversight. The model proposed in this report is designed for public health benefit and does not introduce significant dual-use risks. However, several considerations are worth noting explicitly:

Misuse of aggregated signal data. A system that aggregates weak clinical signals across geographic areas and time could, in principle, be misused to monitor population health patterns without appropriate governance. This risk is mitigated by the design principle of minimal data collection, purpose-specific use, and mandatory human oversight of all aggregated outputs.

Surveillance overreach. There is a risk that a well-functioning sentinel layer could be expanded beyond its intended scope, for instance toward routine population health monitoring unrelated to biosecurity objectives. This risk is addressed through the Tier 1/Tier 2 distinction: Tier 1 captures only a narrow and predefined set of anomaly categories; Tier 2 activation requires explicit public health authority and defined conditions.

Privacy and data protection. Oral health data contains sensitive personal health information. Any implementation must comply with GDPR and applicable national data protection frameworks, with clear boundaries on data access, retention, and secondary use.

Inappropriate automation of clinical escalation. If AI-assisted pattern detection were used to trigger clinical or public health actions without human review, there is a risk of false alerts driving inappropriate responses. This is mitigated by the design principle that all escalation logic remains under public health oversight and is never fully automated.

C. Responsible Disclosure

This report does not present vulnerability disclosures in a conventional security sense. The proposed system does not involve new software, novel biological methods, or the identification of specific weaknesses in existing systems. Its contribution is a policy and systems design argument. If implemented as a pilot, governance protocols for data access, activation authority, and escalation would need to be established in advance and reviewed by relevant public health authorities before operation.

D. Ethical Considerations

The proposal is grounded in ethical proportionality. It is designed to avoid unnecessary expansion of surveillance, protect clinician autonomy, preserve the primacy of patient-centered care, and ensure that any public health benefit is proportionate to governance requirements. The model explicitly avoids diagnostic overreach, continuous surveillance, and extraction of data without clear benefit to participating clinicians or patients.

E. Suggestions for Future Improvements

- Empirical pilot in a bounded public oral health setting to test feasibility and signal quality.
- Development of a validated minimal anomaly taxonomy through participatory design with oral health professionals.
- Integration design study with existing oral health information systems to test technical feasibility of embedded structured capture.
- Governance framework development in collaboration with public health authorities to define Tier 2 activation criteria and data use boundaries.
- Systematic review of existing literature on oral manifestations of emerging infectious diseases to refine the signal taxonomy.
- Extension of the neglected infrastructure concept to other clinical settings — identifying which other routine service environments contain underused surveillance-relevant weak signal.