Youth Innovation Journal
YIJ–2026–0006S

YIJ
YOUTH INNOVATION JOURNAL

VITALIS: A Patient-Owned Agentic Framework for Emergency Detection and Governed Clinical Disclosure

*Authors contributed equally

Anjali Sharma-Tiwari[1*]
Priyank Sharma-Tiwari[1*]
Raeyaan Valleru Muppaneni[2*]
Rishon Valleru Muppaneni[2*]
Rohan Kalavar[3*]
Nirvik Kasula[4*]

Joshua Amaladass[5]
Alex Colin[5]

[1]Basis Independent Silicon Valley, San Jose, California
[2]Irvington High School, Fremont, California
[3]California High School, San Ramon, California
[4]Dougherty Valley High School, San Ramon, California
[5]Robolabs
Dublin, California

Received: 07/13/26
Accepted: 07/15/26
Published: 07/31/26
Corresponding:
innovation@robolabs.org

***Abstract*— Emergency response for a person living alone with chronic illness rarely fails at detection. Consumer wearables already detect falls and irregular rhythms, place emergency calls without user input, and, where users opt in, deliver a static medical profile to dispatchers. What they do not establish is why an alert fired. Detection operates as a binary trigger evaluated against population-calibrated thresholds, with a cancellation window as its only correction, and the resulting false-alarm burden is measurable in dispatch centers. Personal baselines narrow the problem without dissolving it, since a baseline-corrected detector can establish that a signal moved but not what moved it. Meanwhile prehospital clinicians report persistent difficulty accessing patient health information, and community responders to automated alerts name missing information about the cause of the alert among their principal concerns. Among the deployed systems and literature reviewed for this proposal, none interrogates an ambiguous signal before acting on it, and none delivers a governed clinical account to whoever responds.**

This paper presents VITALIS (Vitals Interpretation, Triage, Agentic Local Intelligence System), a framework proposed for that gap. Instruments the person already owns report to an agent hosted on patient-owned hardware, a home hub with a smartphone as its mobile extension, which maintains a contextually conditioned characterization of that individual's ordinary physiology. Deviations are banded by magnitude and source confidence: expected variation is ignored, unambiguous time-critical signals escalate at once, and the ambiguous middle enters a resolution stage before any response is selected. That stage is the distinguishing contribution. It queries the wearer, requests a confirmatory reading from a higher-confidence instrument, and extends the observation window, bounded by a time budget scaled inversely to assessed severity, with unresolved states escalating rather than resting. Response then proceeds through a graded ladder in which each disclosure is scoped by the recipient's role, limited in time, and recorded, extending established break-the-glass discipline past the institutional boundary at which it currently stops. The agent reports observations and their provenance rather than clinical conclusions, leaving diagnosis with the clinician. VITALIS is presented as an architectural proposal, developed to the point where its architecture, operating logic, and limits can be examined and built upon. If developed as intended, it could give a responder arriving at a residence the account a conscious patient would have supplied.

Keywords: Chronic Disease Management, Wearable Sensors, Patient Monitoring, Emergency Medical Services, Intelligent Agents, Edge Computing, Personal Health Records, Health Information Privacy

I. Introduction

Chronic illness is now the ordinary condition of old age rather than the exception. In 2023, an estimated 76.4% of United States adults reported at least one chronic condition, a figure that rises to 93.0% among adults aged 65 and older, while 51.4% of all adults reported two or more [1]. Among noninstitutionalized adults aged 85 and older,

one of the fastest-growing age groups, 37.3% report four or more of eleven selected chronic conditions [2]. Multimorbidity of this kind produces a distinctive clinical profile: several interacting conditions, several medications, and a personal physiological baseline that may differ substantially from population norms. It also produces a distinctive risk. The more conditions a person manages, the greater the likelihood that an acute event will occur at home, without warning, and without a clinician present to interpret it.

For a meaningful share of this population, that event may occur without anyone present at all. Approximately one quarter of community-dwelling Americans aged 65 and older meet objective criteria for social isolation, and the same analysis identifies living alone, chronic illness, and sensory impairment as mutually reinforcing risk factors [3]. The concern in these cases is not only that help may be delayed, but that the mechanisms by which help is ordinarily summoned assume a person capable of recognizing an emergency, locating a telephone, and describing a situation. Loss of consciousness, stroke, hypoglycemia, and cognitive impairment each violate that assumption, and each is more probable in precisely the population least likely to have a second person nearby.

When help is summoned, the information accompanying the request is typically minimal. A survey of 302 ambulance clinicians found that the great majority experienced difficulty accessing patients' health information, that the difficulty was most acute outside normal hours, and that clinicians believed improved access would lead to more appropriate selection of care pathways. The authors concluded by recommending that patient-held records be explored as a remedy [4]. This is not a marginal inefficiency. Prehospital clinicians make consequential decisions about treatment and conveyance on the basis of what can be established at the scene, and a patient who cannot speak carries no history with them. A more recent scoping review protocol observes that physiological data made available at the call-handling stage would enable more accurate communication to paramedics en route, and notes that research into the integration of wearable devices into emergency medical services remains insufficient [5].

The wearable device literature has developed largely in parallel with this problem rather than in response to it. Reviews of wearable sensing in chronic disease have concentrated on ambulatory monitoring and long-term disease management [6], [7], and consumer devices have been evaluated primarily as instruments of behavior change rather than as components of an emergency response pathway. Where emergency capability has been demonstrated, it has been demonstrated promisingly rather than conclusively. Recent work established proof of principle for device-independent smartwatch detection of circulatory arrest, using induced events in a small volunteer cohort and alerting through a test environment rather than live dispatch, with the authors calling for higher specificity and real-world validation [8]. What is no longer speculative is the basic feasibility of automated detection and automated alerting; what remains open is clinical performance at scale and operation within live services. Deployed consumer products already perform fall and crash detection, place emergency calls without user input, and, where users opt in, transmit a static medical profile to emergency communications centers.

What these systems do not do is establish why an alert fired. Four focus groups conducted with fifteen community responders recruited through a national cardiac arrest network identified missing information about the cause of the alert as one of four principal themes, alongside the difficulty of deciding whether to accept an alert at all [9]. The alert arrives as a signal without an account. This limitation follows directly from the architecture that produces it. Consumer detection operates as a binary trigger evaluated against population-calibrated thresholds, with a cancellation window as the only mechanism for correcting a mistake. The consequences are observable. During the 2022-23 winter season, emergency communications centers serving ski regions reported being overwhelmed by automated calls generated by ordinary falls, one county recording a 22 percent rise in hang-ups and misdialled calls year on year, and the manufacturer publicly acknowledged the spike while declining to say how the feature would be changed [10].

Personalization improves detection but does not resolve the underlying ambiguity. Analyses of commercial wearable data identify heterogeneity among and within individuals as a principal obstacle to scaling detection algorithms across a general population, and show that correcting against an individual baseline measurably improves performance [11]. Yet a system built on personal baselines still reports only that a signal has deviated. In a prospective cohort of 3,318 participants, of whom 84 acquired a confirmed infection, a smartwatch alerting system generated alerts for pre-symptomatic or asymptomatic infection in 80% of infected individuals, a median of three days before symptom onset. The same system also generated alerts in response to stress, alcohol consumption, and travel [12]. Deviation from baseline is informative about whether something has changed. It is not informative about what changed, and the distinction is decisive when the available responses range from doing nothing to dispatching an ambulance.

A parallel gap exists on the disclosure side. Emergency access to protected clinical information is a recognized concept, but the standards supply primitives rather than a policy. The HL7 FHIR specification defines how a break-the-glass request is represented and how the resulting access is recorded in an audit event, while stating explicitly that it defines no policy or protocol around such requests and leaving authorization, consent checking, and access control to implementers [13]. Emergency access procedures, audit controls, and minimum-necessary scope are separately required of covered entities under federal health information regulation [14]. Because the policy layer is left open, break-the-glass is realized in practice inside individual electronic health record systems rather than as an interoperable capability. The building blocks for governed emergency disclosure therefore exist, but the mechanism

assembled from them terminates at the institutional boundary and does not reach the responder arriving at a residence.

These gaps are addressed here in light of a shift in what computation on personal devices can be asked to do. An earlier generation of work established that physiological sensing could be miniaturized, embedded in consumer hardware, and worn continuously. A subsequent generation established that models could interpret such signals. What has changed more recently is the emergence of systems that combine a model with planning, persistent memory, and the ability to invoke external tools, so that reasoning is coupled to action rather than terminating in a classification [15]. Reviews of these systems in biomedical settings describe agents that maintain contextual memory across interactions, invoke external services, and execute goal-directed behavior, while cautioning that full autonomy in diagnosis or treatment planning is neither ethically acceptable nor technically safe, and that human oversight must be designed in rather than added afterward [16]. In parallel, work on edge machine learning for wearable biosignals establishes that inference can be placed on the wearable itself or on a paired smartphone, with cloud services reserved for non-time-critical functions, making local processing an architectural choice governed by latency, memory, and energy budgets rather than a technical impossibility [17]. The relevant opportunity is therefore not additional sensing hardware. It is the application of this class of system to the sensing hardware that patients already own.

This paper proposes VITALIS (Vitals Interpretation, Triage, Agentic Local Intelligence System), a conceptual framework for a patient-owned agent that mediates between continuous physiological sensing and emergency response. The framework makes three contributions. First, it inserts an ambiguity resolution stage between detection and escalation, in which the agent actively reduces uncertainty by querying the wearer, requesting a confirmatory measurement from a higher-confidence instrument, and extending the observation window before selecting a response. Second, it replaces binary alerting with a graded escalation ladder in which the severity of the response and the scope of information disclosed increase together. Third, it specifies emergency disclosure as a patient-governed protocol, extending the established break-the-glass model beyond the institutional boundary so that a scoped, time-limited, and auditable clinical summary reaches whoever is responding. The agent is designed to report observations and provenance rather than clinical conclusions, preserving diagnosis as a clinician function.

VITALIS is presented as an architectural proposal. Its components draw on capabilities that exist today in deployed products or in demonstrated prototypes, and the contribution lies in their composition and governance rather than in any single element awaiting invention. The remainder of this paper describes the proposed architecture and operating logic, examines anticipated outcomes and limitations, addresses safety and governance considerations, and identifies directions for further development.

II. Methodology

A. Design Constraints and Scope

VITALIS is presented as an architectural proposal rather than an implementation report, and is described in the present tense throughout. Its components draw on capabilities that exist today in deployed products or in demonstrated prototypes: local inference on consumer hardware, interoperability standards that supply the necessary primitives, and commercially available sensing. The contribution lies in how these are composed and governed rather than in any single element awaiting invention.

The framework targets adults managing one or more chronic conditions, with older adults living alone as the primary case. It assumes commercially available sensing hardware that the person already owns or can obtain through existing clinical channels. It proposes no new sensor.

Four constraints govern every subsequent design decision.

Asymmetry of error. A false alarm imposes cost in money, attention, and responder time. A missed event may cost a life. These are not comparable, and the system is therefore designed so that any state it cannot resolve escalates rather than resting.

Locality. The patient's record and the reasoning over it reside on hardware the patient owns. No third-party service is the system of record. Remote services may be used for non-time-critical functions such as model updates, a division consistent with tiered edge architectures described in the wearable machine learning literature [17].

Observation, not conclusion. The system reports measurements, their provenance, and their relationship to the individual's own baseline. It does not produce diagnoses, does not name conditions, and does not recommend treatment. This constraint is not a limitation reluctantly accepted but a design commitment, and it reflects the position in the biomedical agent literature that full autonomy in diagnosis or treatment planning is neither ethically acceptable nor technically safe, and that human oversight must be architectural rather than supplementary [16].

Graceful degradation. Every component may fail. The design specifies behavior under the loss of each, with conventional device-initiated emergency calling as the final fallback. The system is never the only path to help.

Explicitly out of scope: sensor design and validation, detection algorithm performance, regulatory clearance pathways, and reimbursement.

B. System Architecture

The architecture comprises three tiers separated by a trust boundary (Fig. 1).

The sensing tier consists of heterogeneous devices already in the person's possession: a wrist-worn wearable, a continuous glucose monitor where clinically indicated, an

automated blood pressure cuff, and smartphone sensors. These devices are treated as interchangeable sources conforming to a common interface, described in Section II-C.

The agent tier holds the personal baseline model, the assessment logic, the encrypted health record, and the disclosure control function. It runs across two hosts. A small always-powered home hub serves as the primary, providing continuous availability, mains power, and a fixed network connection. A smartphone serves as the mobile extension, maintaining coverage away from the residence. Either host can carry the full assessment load independently.

The two-host arrangement responds to a specific failure mode. A person living alone requires continuous coverage, and a smartphone is routinely left in another room, allowed to discharge, or powered off. A hub alone cannot follow the person outdoors. Neither host is sufficient; together they provide continuity across the situations in which the target population actually lives.

The recipient tier comprises the parties an escalation may reach: the patient, a caregiver or neighbor, family members, a clinician, and emergency services. Every path from the agent tier to the recipient tier passes through disclosure control. No sensing device communicates outward independently.

The wearable occupies a distinct position within the sensing tier. It is the only component in continuous contact with the body, which makes it both the continuous sensing channel and the sole means by which the system can address the wearer and register the absence of a reply. Its sensing complement is limited to what wrist-worn hardware can deliver passively: photoplethysmographic heart rate, single-lead electrocardiography on demand, skin temperature, and inertial motion. It carries no glucose sensor, non-invasive wrist glucose sensing having no validated implementation, and it does not estimate blood pressure optically, since cuffless estimation is not clinically validated for this purpose. Cleared wrist-worn blood pressure instruments do exist, but they inflate a cuff built into the band and require a deliberate measurement, which places them among the confirmatory sources described in Section II-C rather than in the continuous sensing channel.

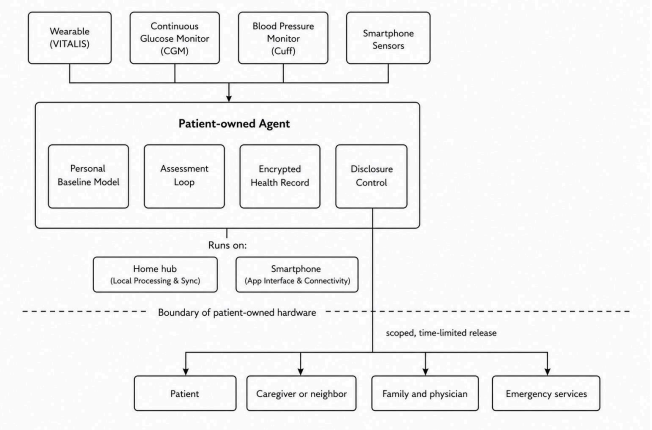

Fig. 1. System architecture. Heterogeneous sensing devices feed a patient-owned agent distributed across a home hub and a smartphone. All outward communication passes through disclosure control at the trust boundary.

C. Input Model and Device Requirements

Rather than specifying supported devices, the framework specifies an interface contract. A device is usable if each reading it supplies carries four elements: the measurand and its value in defined units; a timestamp from a trusted clock; an identifier for the source device and its class; and a quality indicator together with the known accuracy characteristics of that device class.

The fourth element governs how the reading is used. A reading from a validated clinical instrument and an estimate derived from wrist optics are not equivalent evidence, and the framework does not treat them as such. Each reading enters assessment weighted by the confidence characteristics of its source class (Table 1).

This weighting has a direct operational consequence. When a low-confidence source reports a deviation, one available response is not to escalate but to obtain a higher-confidence reading of the same measurand. Confidence weighting is therefore not merely an accounting convenience; it defines part of the resolution repertoire described in Section II-E.

Capability scales with what the person owns. The floor is a wearable alone, supporting cardiac rhythm, motion, and thermal assessment. A continuous glucose monitor extends coverage to metabolic events. A blood pressure cuff extends it to hemodynamic events and supplies a high-confidence confirmatory channel. The framework degrades in capability rather than failing when instruments are absent, and states its capability at each tier rather than claiming universal coverage.

One constraint is not technical. Access to real-time data from commercial sensing platforms is governed by vendor interface policy, and some data is not exportable in real time at all. This is a business and regulatory obstacle rather than a design one, and it is noted here as a genuine barrier to implementation rather than resolved by assumption.

Table 1: Input classes and confidence characteristics

Source class	**Representative measurands**	**Confidence**
Validated clinical instrument	Blood pressure, blood glucose	High
Regulated instrument, continuous or on demand	Interstitial glucose, single-lead ECG	Moderate to high
Consumer optical and inertial	Heart rate, motion, skin temperature	Moderate, context-dependent

Source class	Representative measurands	Confidence
Contextual and self-reported	Activity state, wearer response	Corroborating only

D. Baseline Characterization

The agent maintains a statistical characterization of the individual's ordinary physiology, conditioned on time of day, activity state, and sleep or wake status. An initial learning period establishes it; continuous updating maintains it. That updating must separate two kinds of change: gradual drift over weeks, consistent with normal variation or disease progression, revises the baseline, while acute deviation from it enters assessment. Conflating the two yields either a baseline that chases a deteriorating patient or one that flags ordinary aging.

The baseline describes rather than judges: it characterizes what is usual for this person, not what is healthy for them, preserving the observation-not-conclusion constraint where it is easiest to violate. The rationale is empirical. Heterogeneity among and within individuals is the principal obstacle to scaling detection across a population, and correcting against an individual baseline measurably improves performance [11].

E. Assessment Loop

The assessment loop is the framework's central mechanism (Fig. 2). It replaces the threshold-and-trigger pattern with an explicit stage between detection and escalation.

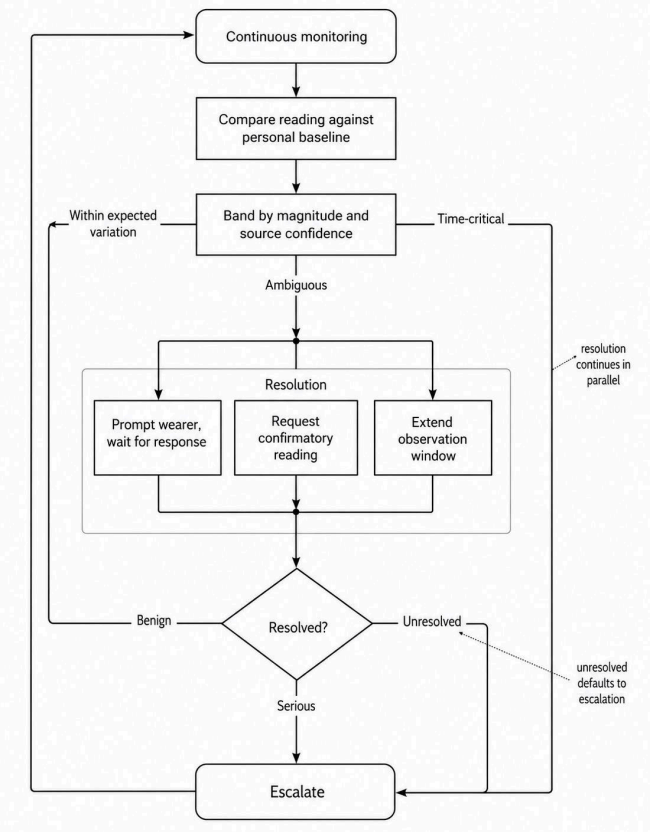

Fig. 2. Assessment loop. Deviations are banded by magnitude and source confidence. Expected variation returns to monitoring, time-critical signals escalate immediately with resolution continuing in parallel, and the ambiguous middle enters a bounded resolution stage. Both resolved-serious and unresolved outcomes escalate; only a resolved-benign outcome returns to monitoring.

Comparison. Incoming readings are compared continuously against the contextually conditioned baseline.

Banding. A deviation is assigned to one of three bands according to its magnitude and the confidence of its source.

Band 1, within expected variation. No action. Monitoring continues.

Band 2, ambiguous. The deviation is real but its cause is undetermined. Resolution is entered before any escalation.

Band 3, unambiguous and time-critical. Signals consistent with catastrophic events, including the absence of detectable circulation, escalate immediately. Resolution then proceeds in parallel, not to gate the alert but to enrich the context accompanying it.

The three-band structure answers the principal objection to this design, which is that gathering more information delays care. Resolution gates escalation only in the ambiguous middle band. At the extremes the loop imposes no delay: Band 1 requires no response, and Band 3 escalates at once. The delay is incurred precisely where the alternative architecture is guessing.

Resolution. Within Band 2, the agent takes one or more actions to reduce uncertainty (Table 2). Each is bounded by a time budget that scales inversely with assessed severity, from seconds at high severity to minutes at low. Expiry of the budget is itself an escalation trigger.

Outcome. Resolution terminates in one of three states. Resolved as benign returns to monitoring, with the episode logged. Resolved as serious escalates. Unresolved escalates.

Two properties of this stage warrant emphasis. First, a wearer's response is an input weighted like any other, not an authoritative verdict. A delayed, incoherent, or internally inconsistent reply is informative and does not resolve toward benign; confusion is a symptom in several of the conditions this framework addresses. Second, non-response never resolves toward benign under any configuration. Silence is the condition the framework exists to interpret, and treating it as reassurance would invert the system's purpose.

Table 2: Resolution actions

Action	Reduces uncertainty about
Query the wearer and await response	Consciousness, orientation, distress

Action	Reduces uncertainty about
Request confirmatory reading from a higher-confidence source	Measurement validity
Extend the observation window	Whether the deviation is transient
Cross-check independent channels	Sensor fault versus physiological event

F. Escalation Ladder

Escalation proceeds through four tiers in which the magnitude of the response and the scope of information disclosed increase together (Fig. 3).

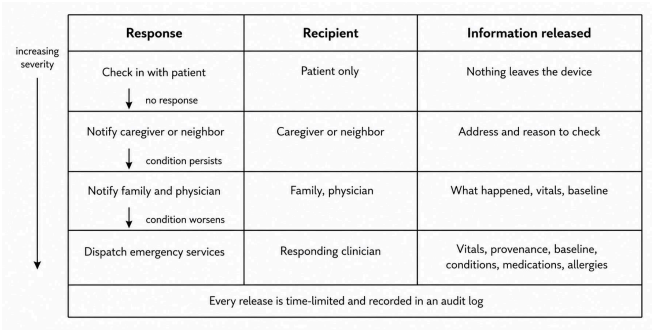

Fig. 3. Escalation ladder, recipients, and corresponding disclosure scope. Response magnitude increases across four tiers, and what each recipient receives is scoped to their role rather than to severity alone. Every release is time-limited and recorded in an audit log retained by the patient.

Tier 1 is a check-in with the patient. Nothing leaves the device. Tier 2 notifies a caregiver or nearby contact. Tier 3 notifies family and, where configured, a clinician. Tier 4 requests emergency medical services.

Advancement is triggered by non-response at the current tier, persistence of the deviation, worsening of the signal, or expiry of the tier's time budget. Tiers may also be entered directly: a Band 3 deviation enters at Tier 4 without traversing the ladder.

Tier membership and thresholds are configured by the patient in advance, not during an event. The patient specifies who occupies each tier, which conditions warrant which response, and whether particular tiers may be skipped. Configuration during an emergency is not assumed to be possible.

Cancellation is asymmetric by tier. At Tiers 1 through 3 the patient may cancel and the system returns to monitoring with the episode logged. At Tier 4 the system does not unilaterally retract a request for emergency services; cancellation follows the receiving dispatcher's protocol. Whether an ambulance already dispatched is still needed is a clinical determination, and the framework does not claim authority over it.

G. Disclosure Protocol

The patient's clinical record is held encrypted on patient-owned hardware. Emergency disclosure is governed rather than open: each release is scoped to a defined recipient, limited in time, and recorded.

This builds on an established idea while supplying what the standards deliberately leave open. The HL7 FHIR specification provides the representational building blocks for break-the-glass access, including a means of marking such a request and an audit event for recording it, while stating that it defines no policy or protocol governing those requests [13]. Emergency access procedures, audit controls, and minimum-necessary scope are separately required of covered entities under federal health information regulation [14], though those obligations attach to covered entities and business associates rather than to every party in this architecture, so the framework adopts them as design requirements rather than inheriting them as legal duties. What is missing is not the primitives but the policy layer and its confinement to institutional systems. The proposal here is to supply that layer and extend it outward, so that a scoped and auditable release can reach the responder at the residence.

The released packet contains observations, their provenance, and their relationship to the individual's baseline. It states, for example, that heart rate was measured at 38 beats per minute at a given time by a named device class, that the individual's ninety-day resting baseline is 62, and that two check-ins received no response. It does not state that a cardiac event is probable. Interpretation is the clinician's function, and the packet is constructed so that the clinician receives what they need to perform it.

Each release carries an expiry, terminating at incident close or at a fixed timeout, whichever is earlier. An audit record retained by the patient logs what was released, to whom, at what time, and under which tier.

Recipient authentication is an unresolved problem. Verifying that a party requesting or receiving a release is genuinely the responding clinician, without introducing a step that delays care, is not solved by this framework. It is identified here as a requirement for any implementation and revisited in Section V.

H. Interaction Design

The check-in is the mechanism on which the assessment loop depends, and its design cannot assume a capable respondent.

It proceeds across three modalities of escalating salience: haptic alert at the wrist, audible prompt, and visual prompt on the wearable display (Fig. 4). Haptic delivery is primary rather than supplementary, since hearing loss is common in the target population and an audio-only prompt would fail silently for a substantial fraction of intended users. Response may be given by touch, by voice, or by a deliberate gesture, so that no single impairment prevents a reply.

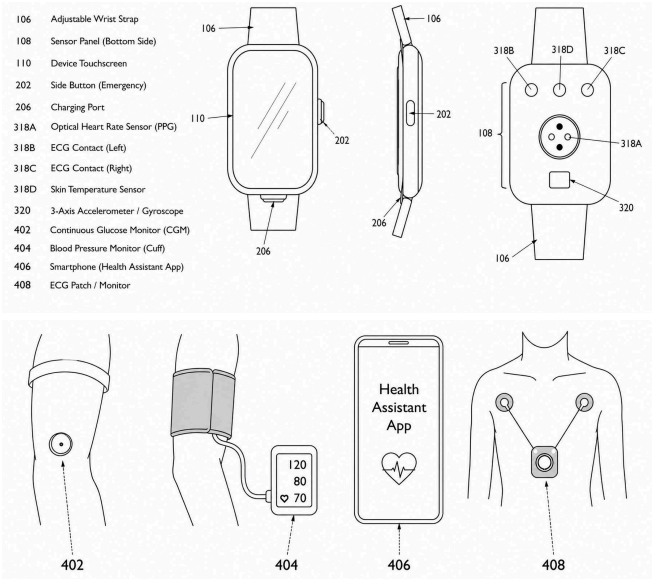

Fig. 4. Reference sensing endpoint and supporting instruments. (a) Wearable exterior and skin-contact layout showing the sensing complement available from wrist-worn hardware. (b) External instruments contributing to assessment, all reporting to the agent over Bluetooth Low Energy.

Response interpretation accounts for the population. Reduced dexterity, disorientation, and cognitive impairment are prevalent, and a coherent verbal answer cannot be treated as the expected case. The framework therefore accepts weak evidence of responsiveness, including a purposeful gesture or a partial reply, as evidence of consciousness while continuing assessment on the physiological signals.

Cancellation at Tiers 1 through 3 is deliberately easy. This is a safety property rather than a convenience. A system that is difficult to silence during false alarms will be disabled by the people it protects, and a disabled system protects no one. The design assumes false alarms will occur, and treats their handling as a functional requirement rather than a failure to be minimized out of existence.

III. Anticipated Outcomes and Limitations

VITALIS is an architectural proposal, developed far enough that its operating logic, its dependencies, and the conditions under which it would fail can be examined and argued with. This section sets out the outcomes the proposed framework is intended to produce, each stated conditionally, and then the constraints that would bound them. Subsections A through D describe the anticipated outcomes; E through I describe the limitations. Both are given full weight. A design at this stage is most useful when the boundary of what it could do is drawn as carefully as the capability itself, and it is that pairing which makes the proposal testable by whoever builds on it.

A. Clinical Context at the Point of Response

The best supported of the proposed outcomes is the delivery of clinical context to whoever responds. The need is documented from both ends of the encounter. A survey of 302 prehospital clinicians found that the great majority experienced difficulty accessing patients' health information, that the difficulty was most acute outside normal hours, and that improved access was judged likely to produce more appropriate selection of care pathways; the authors' own recommendation was that patient-held records be explored [4]. From the responder side, focus groups with community responders to smartwatch-initiated alerts identified missing information about the cause of the alert as one of four principal themes [9]. A recent scoping review protocol makes the same observation from the service perspective, noting that physiological data available at the call-handling stage would enable more accurate communication to paramedics en route [5].

The framework's disclosure protocol is intended to close that gap by delivering, at the moment of escalation, a scoped summary containing the measurements that prompted the response, their provenance, and their relationship to the individual's own baseline. If realized as proposed, the outcome would be that a responder arriving at a residence begins with the information a conscious patient would otherwise have supplied, in circumstances where the patient cannot supply it.

This outcome is stated with more confidence than the others because the components it depends on already exist. Representational primitives for break-the-glass access and audit recording are specified [13], and emergency access, audit controls, and minimum-necessary scope are established requirements for covered entities [14]. What is proposed is not a new mechanism but the policy layer the standards leave open, applied past the institutional boundary at which existing implementations stop.

The same mechanism serves recipients who are not clinicians, and what each receives is scoped by their role and not only by the severity that triggered the release. A neighbor asked to look in on someone needs an address and a reason to go, not a medication list. A family member needs to know what happened and what was done. A responding clinician needs the measurements, their provenance, and the baseline against which they are abnormal. Scoping by recipient role as well as by tier keeps the minimum-necessary principle [14] intact at every rung of the ladder rather than only at its top, and it means the escalation sequence functions as a routing structure for information as much as for response: the framework decides not only how much to do but who needs to know what in order to do it.

B. Proportionate Response Through Ambiguity Resolution

The distinguishing outcome, and the framework's principal novel contribution, concerns what happens between detection and escalation. Current systems resolve a deviation into an alert or into silence, and the cost of that binary is visible in operation: during the 2022-23 winter season, emergency communications centers serving ski regions reported being overwhelmed by automated calls generated by ordinary falls, with one county recording a 22 percent year-on-year rise in hang-ups and misdialled calls [10].

Personalization narrows the problem without

dissolving it. Correcting against an individual baseline measurably improves detection performance, and heterogeneity among and within individuals is identified as the principal obstacle to scaling detection across a general population [11]. Yet a baseline-corrected detector still reports only that a signal has moved. In a cohort of 3,318 participants, a smartwatch alerting system generated alerts for pre-symptomatic or asymptomatic infection in 80% of 84 infected individuals a median of three days before symptom onset, while also generating alerts in response to stress, alcohol consumption, and travel [12]. The system could establish that something had changed and not what.

The anticipated outcome of the resolution stage is that a deviation of undetermined cause is interrogated before it is acted upon, and that the response selected is proportionate to what the interrogation establishes. Where resolution returns a benign explanation, no one is called. Where it returns a serious one, or returns nothing before its time budget expires, escalation proceeds. The intended effect is not fewer alerts but better-matched ones: a check-in where a check-in suffices, and an ambulance where an ambulance is warranted.

A distinction has to be drawn here, because deployed systems already pause before escalating. Fall detection presents a countdown and an opportunity to cancel. That is a cancellation window, not a resolution stage. Its timeout is fixed rather than scaled to severity, its only input is whether the user intervened, and a user who does not intervene is treated identically whether they are unconscious, asleep, or merely not looking at their wrist. The stage proposed here differs in three respects: it gathers evidence from channels other than the user, it weighs whatever the user provides against the physiological picture rather than accepting it as an override, and it terminates in an assessment rather than in the expiry of a timer. The difference is between offering a chance to cancel and attempting to establish what happened.

One asymmetry within that stage should be stated plainly. The confirmatory channels available to it are not equally accessible in all conditions. Continuous instruments such as an interstitial glucose monitor can be interrogated without the wearer's participation, whereas a cuff measurement or an on-demand electrocardiogram requires a deliberate action. The confirmatory branch is therefore strongest in the ambiguous-but-conscious case and weakest precisely where the wearer is unresponsive, which is the case of greatest concern. In that circumstance the stage falls back to passive channels and to the observation window, and the absence of a reply becomes itself the dominant evidence.

This outcome is stated with considerably more caution than the preceding one, because the mechanism it depends on has not been demonstrated. That limit is set out first among the limitations below, in Section III-E, and every capability attributed to the resolution stage here is conditional on it.

C. Continuous Attention Without Continuous Cost

Continuous, informed attention to a chronically ill person is not a new capability. It exists, in the form of private nursing care and staffed residential facilities, and it works. What distinguishes it is not its difficulty but its distribution: it is allocated by ability to pay or by placement into an institution, and the population that most needs it is disproportionately outside both. Nearly all adults aged 65 and older live with at least one chronic condition and a substantial minority of those over 85 live with four or more [1], [2], while roughly one quarter of community-dwelling older adults meet objective criteria for social isolation [3].

Two properties of the framework bear on that distribution. The first is that it proposes no new sensing hardware. The interface contract of Section II-C admits instruments the person already owns or obtains through existing clinical channels, which means the marginal cost of the sensing tier is zero for a household already equipped and incremental for one that is not. The second is that the capability is expressed in software running on general-purpose hardware, which does not need to be staffed, scheduled, or physically present.

The anticipated outcome is therefore that a form of continuous attention which is currently a staffing decision or a placement decision could instead become a deployment decision, and that its availability would not depend on being able to purchase a person's time.

One qualification belongs with this outcome and should not be left to a reader to supply. Attention is not care. The framework watches, interprets, and summons; it does not treat, assist, or keep company, and the isolation described in the literature it cites is a condition it responds to rather than one it remedies [3]. There is a foreseeable misuse in which the presence of such a system is taken as grounds for reducing human contact with an older adult living alone. That would be a harm produced by the framework rather than prevented by it, and nothing in the design guards against it.

D. Patient-Held Infrastructure

The decision to place the record and the reasoning over it on patient-owned hardware is described in Section II-A as a design constraint. Four outcomes follow from it, and only one is about privacy.

The first is availability. A system whose failure mode is that no help arrives cannot sensibly inherit the availability of a network connection and a remote provider. Under the proposed architecture, loss of connectivity degrades reach, since remote contacts cannot be notified, but it does not stop sensing, assessment, or local alerting. A cloud-dependent equivalent loses all four at once, and it loses them in precisely the circumstances, rural connectivity and infrastructure disruption, where the population is least well served to begin with.

The second is continuity. Consumer health services are discontinued, acquired, and repositioned, and a monitoring capability that depends on a specific company's continued interest in offering it is a capability with an expiry date that its user does not control. A framework specified as

an architecture rather than a product can be implemented by more than one party, and a household that has adopted it does not lose it when a vendor changes direction.

The third is privacy. The largest concentration of sensitive data sits where the patient controls it rather than in infrastructure that aggregates many users, so a single compromise does not expose a population. The disclosure protocol of Section II-G then governs each departure from that store individually.

The fourth is authority over policy. Because escalation tiers and thresholds are configured by the patient in advance, the question of who is called, and under what conditions, is settled by the person the system serves rather than by a default chosen elsewhere.

This arrangement is feasible rather than aspirational. Work on edge machine learning for wearable biosignals establishes that inference can be placed on the wearable or a paired device, with remote services reserved for functions that tolerate latency, making local processing a matter of latency, memory, and energy budgets rather than an impossibility [17].

The cost is real and falls on the household. Patient-held infrastructure means hardware to acquire, configure, keep powered, update, and eventually replace, and the population this framework addresses is the one least likely to do that unassisted. A caregiver, family member, or clinician sits in the setup and maintenance path by necessity, which narrows the outcome described above: what becomes available without purchasing a person's time still requires a person's involvement at the outset.

E. The Resolution Stage Has Not Been Demonstrated

The most important limitation concerns the mechanism on which the framework's novelty rests. No published work demonstrates that an active resolution stage reduces inappropriate escalation without delaying appropriate escalation, because no such stage has been built and evaluated. The supporting literature establishes the components of the argument and not the argument itself: that binary triggering produces a real false-alarm burden [10], that population-calibrated thresholds scale poorly across individuals [11], and that baseline-corrected detection still cannot distinguish a medical event from a benign one [12]. That a bounded interrogation would improve on this is an expectation derived from the structure of the problem, not a measured result.

Two specific quantities are unknown and would decide whether the design works. The first is the distribution of deviations across the three bands defined in Section II-E. If the ambiguous band is narrow, the resolution stage rarely engages and contributes little. If it is wide, the stage engages constantly and the system becomes intrusive. Neither the width of that band nor its dependence on the individual has been characterized. The second is the relationship between resolution latency and clinical outcome. The design assumes that a bounded delay in the ambiguous band is preferable to guessing, and bounds that delay by a budget scaled to assessed severity. Whether the budgets proposed are clinically acceptable is a question for outcome data that does not yet exist.

Until both are established, the outcome described in Section III-B remains an expectation grounded in the structure of the problem rather than a capability that can be assumed.

F. Data Access Is Constrained by Policy Rather Than Design

The framework specifies an interface contract rather than a device list, which resolves the technical side of heterogeneous input. It does not resolve the commercial side. Access to real-time data from consumer sensing platforms is governed by vendor interface policy, and some measurements are not exportable in real time at any price. A framework that depends on ingesting from instruments the patient already owns therefore depends on the willingness of the companies that made them.

This limitation is qualitatively different from the others, in that no amount of design work removes it. It is noted here rather than resolved, and it is worth stating that a patient-owned architecture is on the less favored side of the incentive: the platforms best positioned to supply the data are the platforms whose own products the framework would displace as the system of record.

G. The Framework Assumes a Wearer, and Sometimes a Responsive One

Two assumptions about the person run underneath the design.

The first is that the device is being worn. Continuous physiological monitoring produces nothing during the intervals when a wearable is charging, has been removed for bathing, or has been forgotten, and adherence to wearable use declines over time in exactly the population this framework addresses. The design's fallback during such gaps is the absence of monitoring rather than a degraded form of it, and an event occurring in that window is invisible to the system.

That gap has an engineering cause worth naming. Continuous multi-modal sensing, on-device inference, and radio activity are the three largest consumers in a wrist-worn power budget, and a design that increases all three shortens the interval between charges. The relationship is therefore adversarial: the more capable the sensing, the more often the device is off the wrist and the wider the blind interval. Duty-cycled sensing, in which sampling rates and inference frequency rise only once a deviation is detected, is the conventional means of managing this and is assumed here rather than specified. The resulting power envelope, and the detection sensitivity that survives duty cycling, are among the first quantities any prototype would have to establish.

The second is subtler. The resolution stage draws part of its information from the wearer's response, and Section II-H specifies that weak evidence of responsiveness is accepted while assessment continues on the physiological signals. This is a deliberate accommodation of hearing loss, reduced dexterity, and cognitive impairment, all prevalent in the intended population [3]. But the accommodation has a

cost that should be stated: a wearer experiencing hypoglycemia, stroke, or delirium may produce a response that appears adequate while their condition is deteriorating. The design mitigates this by never allowing a response to override the physiological picture, but mitigation is not elimination, and the failure mode is real.

H. Recipient Authentication Remains Unsolved

Verifying that a party receiving a clinical release is genuinely the responding clinician, and doing so without introducing a verification step that costs time in an emergency, is not solved by this framework.

The difficulty is structural. Strong authentication assumes an enrolled party with credentials that can be checked against a directory, and emergency response frequently involves whoever arrives. Weak authentication makes the disclosure protocol a liability rather than a safeguard, since a system that releases a clinical dossier to any requester is worse than one that releases nothing. Existing break-the-glass implementations avoid the problem by operating inside an institution where every user is enrolled [13], which is precisely the boundary this framework proposes to cross. Any implementation would need to answer this, and the answer is more likely to come from emergency communications infrastructure than from the framework itself.

I. What Remains to Be Tested

Every outcome above is inferred from the architecture set out in Section II and from the prior work cited in Section I rather than measured. Establishing whether VITALIS would perform as intended would require a staged program, each stage testing a specific claim made in this section rather than the concept as a whole.

The first stage is retrospective and requires no hardware. Archived continuous physiological datasets can be replayed through the banding logic of Section II-E to characterize how deviations distribute across the three bands, how that distribution varies between individuals, and how often a benign explanation would have been available from a second channel. This addresses the first unknown named in Section III-E and is the cheapest useful experiment available.

The second stage is a human factors study of the check-in with the intended population rather than with convenience participants. Whether a multi-modal prompt reliably elicits an interpretable response from older adults with hearing loss or mild cognitive impairment is an empirical question, and the accommodation described in Section II-H is currently a design intention rather than a validated interaction.

The third stage is a supervised prospective deployment in which the framework runs end to end but escalation is advisory: proposed responses are logged and reviewed by a clinician rather than acted upon. This would produce the first evidence on resolution latency and on the rate at which the system reaches each tier, without exposing anyone to the consequences of an untested decision.

Only after those would live escalation be appropriate, and then at a scale small enough to be observed. That sequence is the natural continuation of the work presented here.

IV. Safety, Privacy, and Regulatory Posture

A system that reads a person's physiology continuously, holds their clinical history, and can summon an ambulance on their behalf operates inside an existing body of regulation and professional practice. That body is not a constraint applied to the concept after the fact; it shaped several of the choices described in Section II, most directly the decision that the agent reports observations rather than conclusions. This section states where the proposal sits against those frameworks and the posture it adopts in response.

A. Why the Agent Does Not Diagnose

A system with continuous access to a patient's physiology, medications, and history could in principle produce an interpretation of what is happening to them. VITALIS does not, for reasons that are simultaneously ethical, technical, and regulatory.

The ethical and technical case is made in the biomedical agent literature rather than by this paper. Reviews of large language model agents in biomedicine conclude that full autonomy in diagnosis or treatment planning is neither ethically acceptable nor technically safe, and that human oversight must be architectural rather than supplementary [16]. Related work positions such systems as collaborators that assemble and surface evidence rather than as substitutes for expert judgment [15].

The design implements that conclusion literally. The escalation packet described in Section II-G reports that heart rate was measured at 38 beats per minute at a stated time by a named device class, that the ninety-day resting baseline is 62, and that two check-ins received no response. It does not report that a cardiac event is probable, name a condition, or propose a treatment.

The consequence for responsibility is direct. Diagnosis remains a clinician function, and the clinician receives what they need to perform it rather than a conclusion to accept or overturn. Sensor accuracy remains with the device manufacturers, as it already does. The patient configures their own escalation policy. What remains attributable to the framework is narrower: that the data faithfully represents what the instruments reported, that messages were delivered, and that the audit record is complete. The design intention is a liability posture closer to that of a smoke alarm than of a physician. Whether it would be assigned that posture is a legal question this paper does not attempt to settle, and Section IV-B notes that classification and the liability that follows from it remain unresolved.

B. Regulatory Classification Is Undetermined

Whether a system of this kind is a regulated medical device in the United States would depend on its

claims and functions rather than on its architecture alone, and this paper does not resolve the question.

Two Food and Drug Administration guidance documents issued in January 2026 define the boundaries. The general wellness guidance maintains that low-risk products intended solely for general wellness either fall outside the device definition or are subject to enforcement discretion rather than device requirements, and extends that treatment to certain sensor-based wearables, while noting that products estimating physiological parameters may fall outside the category where their claims or outputs imply medical use [18]. The clinical decision support guidance interprets the statutory criteria excluding software from the device definition, weighting whether a clinician can independently review the basis of what the software presents and observing that the more a system operates as a black box, the more likely it is to be treated as a device [19].

VITALIS sits across this line rather than comfortably on one side. It does not diagnose or recommend treatment, and it presents measurements with their provenance so that the basis of any output is inspectable by the clinician receiving it. But a system that initiates emergency dispatch on the basis of physiological inference is not plausibly a general wellness product, and the resolution logic of Section II-E is an inference about clinical significance even when its output is expressed as an observation.

The proposal does not treat that ambiguity as permission. The choices that keep the agent inspectable and non-diagnostic were made for the reasons in Section IV-A and would remain correct under either classification; they reduce risk rather than evade oversight. Any implementation would require a classification determination before deployment.

C. Data Protection and the Threat Model

Section III-D describes what patient-held infrastructure is intended to achieve. This subsection addresses what it does not remove.

Every departure from the local store is governed by the disclosure protocol of Section II-G, which applies an established discipline to each release: authorization to declare, minimum-necessary scope, time limitation, and mandatory audit [13], [14]. Scoping by recipient role as well as by escalation tier is what makes the minimum-necessary requirement meaningful in practice, since a neighbor and a responding clinician have different legitimate needs and a single graded disclosure would over-serve one of them. Because the discipline applies to every release, including those to family, the framework never operates in a mode where the record is simply open.

Locality relocates risk rather than eliminating it. A home hub is a physical device in an unmonitored residence, and that threat model is not obviously better than a hardened data center operated by a company with a security team. The architecture changes the shape of the exposure rather than its magnitude: a compromise reaches one household instead of a population, but it is also less likely to be noticed. Encryption at rest and the audit record reduce exposure and support accountability without guaranteeing detection, and neither prevents offline extraction following physical compromise. A stronger claim would require tamper resistance and external monitoring. That is a weaker guarantee than a patient-owned architecture might appear to promise.

D. The Design's Governance Posture

Four features follow from the considerations above, each carrying a cost accepted deliberately.

The agent reports observations and never conclusions. The cost is that a responder receives less than the system could in principle provide, and must perform an interpretation the system might have attempted.

Escalation policy and disclosure scope are configured by the patient in advance, so that who is contacted at each tier, and what each of them receives, reflect a decision made by an identifiable person under no time pressure. The cost is that a policy set months earlier may not fit the situation that arises.

Every disclosure is scoped, time-limited, and recorded, which is protective but also accountable, since the audit trail records what was released rather than what was intended. The cost is that the audit record is itself sensitive.

Any state the system cannot resolve escalates. This is a governance position as much as a technical one: the framework fails toward summoning help rather than toward silence, and will therefore sometimes summon help that was not needed. Section II-H accepts that consequence and treats easy cancellation at the lower tiers as a functional requirement, on the reasoning that a system users cannot silence is a system users will disable.

None of this substitutes for the regulatory and ethical review an actual deployment would require, and it is not offered as evidence that such review would be satisfied. It is intended to make the proposal legible to that review: to state in advance what the system will and will not assert, who decides what happens, what may leave the device, and how anyone would later establish what occurred.

V. Future Work

The staged program described in Section III-I sets the order of the work that follows. Each item below is deferred rather than claimed.

The first and most consequential is characterizing the ambiguous band. Section III-E identifies the width of that band, and its variation between individuals, as the quantity that decides whether the resolution stage contributes anything: too narrow and the stage rarely engages, too wide and the system becomes intrusive. This question can be approached now, without hardware, by replaying archived continuous physiological datasets through the banding logic of Section II-E. Two qualifications belong with it. A retrospective study establishes how often a deviation would have been ambiguous, not whether interrogating it would have helped,

so it bounds the problem rather than validating the remedy. And the question is partly independent of this framework, since any system offering graded responses requires the same characterization; progress may therefore come from work unconnected to VITALIS, which would serve the design rather than threaten it.

Whether the check-in works with the intended population is the next question, and it is a human factors question rather than an engineering one. Section II-H specifies multi-modal prompting with haptic delivery as primary, and accepts weak evidence of responsiveness on the reasoning that hearing loss, reduced dexterity, and cognitive impairment are prevalent among the intended users. That is a design intention. Whether such a prompt reliably elicits an interpretable response from an older adult with mild cognitive impairment, and whether an apparently adequate response can be distinguished from a genuinely reassuring one, would require study with those users rather than with convenience participants.

The power envelope is a further line, and it constrains everything above it. Section III-G describes the adversarial relationship between sensing capability and wear time: continuous multi-modal sensing, on-device inference, and radio activity are the largest consumers in a wrist-worn power budget, and increasing them widens the interval during which the device is off the wrist. Duty-cycled sensing is assumed here rather than specified, which leaves two quantities open: the power envelope a duty-cycled implementation would actually achieve, and the detection sensitivity that survives the reduced sampling. Neither can be settled by argument.

Recipient authentication is deferred to infrastructure rather than to this design. Section III-H sets out why the problem is structural: strong authentication assumes an enrolled party with checkable credentials, and emergency response frequently involves whoever arrives. The likely path runs through emergency communications infrastructure, where responding units are already known to a dispatching authority, rather than through anything a patient-held system could establish on its own. Any implementation would need to adopt whatever mechanism that infrastructure provides.

Interoperable access to real-time data from consumer sensing platforms is the one constraint that no amount of design work removes, as Section III-F states. Its resolution would come through vendor policy, through interoperability regulation, or through instrument classes that expose their readings by default. The framework is specified as an interface contract precisely so that it can accommodate whatever access becomes available rather than depending on a particular platform, but a contract cannot compel anyone to meet it.

A classification determination would precede any deployment, as Section IV-B states, and it belongs to the first stage of an implementation program rather than to research. Its outcome would shape the work above: a system classified as a medical device would face validation requirements that a general wellness product would not, and the staged program in Section III-I would expand accordingly.

A final note closes the list. This paper offers no basis for assessing cost, and none should be inferred from it. The sensing tier is designed to carry no marginal cost for a household already equipped, but patient-held infrastructure is not free: a hub must be acquired, powered, maintained, and eventually replaced, and Section III-D notes that the population addressed here is the least likely to do that unassisted. Whether the arrangement is affordable at the scale it would need to reach, and who would bear that cost, are questions for an analysis resting on hardware that has been built and a deployment pattern that has been observed.

VI. Conclusion

VITALIS is an architectural proposal for a patient-owned agent that mediates between continuous physiological sensing and emergency response. Sensing instruments the person already owns report to an agent distributed across a home hub and a smartphone, which maintains a characterization of that individual's ordinary physiology, interrogates deviations of undetermined cause before acting on them, and selects a response from a graded ladder. Where a response requires reaching someone else, a scoped, time-limited, and recorded release delivers to each recipient what their role requires: an address and a reason to a neighbor, an account of what happened to a family member, and measurements with their provenance and baseline comparison to a responding clinician.

The constraints are narrow by choice, and the narrowness is a commitment rather than an omission. The agent reports observations and never conclusions, leaving diagnosis where it belongs. The record and the reasoning over it remain on hardware the patient owns. Escalation policy is configured by the patient in advance rather than by a vendor default. And any state the system cannot resolve escalates, which means it will sometimes summon help that was not needed. Those choices bound what the framework could do and would make it answerable for what it did.

What the paper does not offer should be equally clear. The resolution stage on which the contribution rests has not been built or measured, and Section III names the specific unknowns rather than leaving them for a reader to find: the width of the ambiguous band, the relationship between resolution latency and clinical outcome, the power envelope a duty-cycled implementation would achieve, the reduced accessibility of confirmatory readings when the wearer is unresponsive, and the unsolved problem of authenticating whoever receives a release.

What distinguishes the proposal is the stage it inserts where existing systems have none. A consumer device asks whether to raise an alarm and offers a window in which to cancel it. The framework proposed here asks what is

actually happening, spends a bounded interval finding out, and then asks who needs to know in order to help. That shift, from a device that reports an event toward an agent that establishes one and explains it to whoever answers, is this paper's contribution. Whether the interval it spends is short enough to be safe, and whether the ambiguity it resolves is common enough to be worth resolving, are the questions the next stage of the work must answer.

VII. Acknowledgments

We thank our mentors for their guidance throughout the development of the VITALIS (Vitals Interpretation, Triage, Agentic Local Intelligence System) concept, and for the questions that sharpened it. We are grateful to Robolabs for the collaboration, technical resources, and program support that made this work possible, and to the Innovation Academy for the setting in which it was carried out. We also thank the colleagues and reviewers whose comments improved the design proposal presented here.

References

[1] K. B. Watson, J. L. Wiltz, K. Nhim, R. B. Kaufmann, C. W. Thomas, and K. J. Greenlund, "Trends in multiple chronic conditions among US adults, by life stage, Behavioral Risk Factor Surveillance System, 2013–2023," *Prev. Chronic Dis.*, vol. 22, art. 240539, 2025, doi: 10.5888/pcd22.240539.

[2] Y. Gorina and E. A. Kramarow, "Chronic conditions in adults age 85 and older: United States, 2022–2023," National Center for Health Statistics, Hyattsville, MD, USA, NCHS Health E-Stat 105, Jun. 2025, doi: 10.15620/cdc/174611.

[3] National Academies of Sciences, Engineering, and Medicine, *Social Isolation and Loneliness in Older Adults: Opportunities for the Health Care System*. Washington, DC, USA: National Academies Press, 2020, doi: 10.17226/25663.

[4] O. Zorab, M. Robinson, and R. Endacott, "Are prehospital treatment or conveyance decisions affected by an ambulance crew's ability to access a patient's health information?," *BMC Emerg. Med.*, vol. 15, art. 26, Oct. 2015, doi: 10.1186/s12873-015-0054-1.

[5] R. Alotaibi, A. Alghaith, L. Hughes Noehrer, G. B. Kitchen, and R. Body, "Use of wearable technology in improving emergency care and health outcomes for patients with urgent health complaints: Protocol for a scoping review," *BMJ Open*, vol. 16, no. 3, art. e106396, Mar. 2026, doi: 10.1136/bmjopen-2025-106396.

[6] F. P. Wieringa, N. J. H. Broers, J. P. Kooman, F. M. Van Der Sande, and C. Van Hoof, "Wearable sensors: Can they benefit patients with chronic kidney disease?," *Expert Rev. Med. Devices*, vol. 14, no. 7, pp. 505–519, Jul. 2017, doi: 10.1080/17434440.2017.1342533.

[7] M. A. Kirk, M. Amiri, M. Pirbaglou, and P. Ritvo, "Wearable technology and physical activity behavior change in adults with chronic cardiometabolic disease: A systematic review and meta-analysis," *Amer. J. Health Promot.*, vol. 33, no. 5, pp. 778–791, Jun. 2019, doi: 10.1177/0890117118816278.

[8] W. M. F. van den Beuken *et al.*, "Automated cardiac arrest detection and emergency service alerting using device-independent smartwatch technology: Proof-of-principle," *Resuscitation*, vol. 213, art. 110657, Aug. 2025, doi: 10.1016/j.resuscitation.2025.110657.

[9] M. Eversdijk, B. J. W. van der Eerden, D. L. Willems, W. J. Kop, M. Habibović, and M. A. R. Bak, "Integrating smartwatch-based out-of-hospital cardiac arrest detection into resuscitation systems: A focus group study of community responder perspectives," *Resuscitation Plus*, vol. 25, art. 101027, Sep. 2025, doi: 10.1016/j.resplu.2025.101027.

[10] D. Price, "Apple acknowledges faulty crash detection affecting skiers," *Macworld*, Jan. 16, 2023. [Online]. Available: https://www.macworld.com/article/1473802/apple-acknowledges-faulty-crash-detection-skiers.html

[11] J. Phipps, B. Passage, K. Sel, J. Martinez, M. Saadat, T. Koker, N. Damaso, S. Davis, J. Palmer, K. Claypool, C. Kiley, R. I. Pettigrew, and R. Jafari, "Early adverse physiological event detection using commercial wearables: Challenges and opportunities," *npj Digit. Med.*, vol. 7, art. 136, May 2024, doi: 10.1038/s41746-024-01129-1.

[12] A. Alavi, G. K. Bogu, M. Wang, E. S. Rangan, A. W. Brooks, Q. Wang, E. Higgs, A. Celli, T. Mishra, A. A. Metwally, K. Cha, P. Knowles, A. A. Alavi, R. Bhasin, S. Panchamukhi, D. Celis, T. Aditya, A. Honkala, B. Rolnik, E. Hunting, O. Dagan-Rosenfeld, A. Chauhan, J. W. Li, C. Bejikian, V. Krishnan, L. McGuire, X. Li, A. Bahmani, and M. P. Snyder, "Real-time alerting system for COVID-19 and other stress events using wearable data," *Nature Med.*, vol. 28, no. 1, pp. 175–184, Jan. 2022, doi: 10.1038/s41591-021-01593-2.

[13] HL7 International, "Security Labels," FHIR Specification, Release 5 (v5.0.0). [Online]. Available: https://hl7.org/fhir/security-labels.html

[14] U.S. Department of Health and Human Services, "Security Standards for the Protection of Electronic Protected Health Information," 45 C.F.R. pt. 164, subpt. C, including the emergency access procedure at 164.312(a)(2)(ii) and audit controls at 164.312(b). [Online]. Available: https://www.hhs.gov/hipaa/for-professionals/security/index.html

[15] S. Gao *et al.*, "Empowering biomedical discovery with AI agents," *Cell*, vol. 187, no. 22, pp. 6125–6151, Oct. 2024, doi: 10.1016/j.cell.2024.09.022.

[16] X. Xu and R. Sankar, "Large language model agents for biomedicine: A comprehensive review of methods, evaluations, challenges, and future directions," *Information*, vol. 16, no. 10, art. 894, Oct. 2025, doi: 10.3390/info16100894.

[17] M. Oloko-Oba, E. Esenogho, and K. Aruleba, "From biosignals to bedside: A review of real-time edge machine learning for wearable health monitoring," *Bioengineering*, vol. 13, no. 5, art. 559, May 2026, doi: 10.3390/bioengineering13050559.

[18] U.S. Food and Drug Administration, *General Wellness: Policy for Low Risk Devices*, Guidance for Industry and Food and Drug Administration Staff, Jan. 6, 2026. [Online]. Available: https://www.fda.gov/media/90652/download

[19] U.S. Food and Drug Administration, Clinical Decision Support Software, Guidance for Industry and Food and Drug Administration Staff, Jan. 29, 2026 (superseding the version issued Jan. 6, 2026). [Online]. Available: https://www.fda.gov/media/109618/download