forum acusticum 2026
8-12 September · Graz, Austria

A HUMAN-CENTRED MULTILAYER FRAMEWORK FOR CONTINUOUS ACOUSTIC CHARACTERISATION OF HEALTHCARE ACOUSTIC ENVIRONMENTS

Alejandro Solorio[1], Finn Krewer[2], Eoin King[1]
[1] *The Galway Sound Lab, School of Engineering, University of Galway, Ireland*
[2] *HRB Clinical Research Facility, University Hospital Galway, Ireland*

Abstract

Healthcare acoustic environments (HAEs) are commonly characterised using sound level descriptors (e.g. LAeq, L10, Lmax, etc). These quantify acoustic magnitude and temporal variation but provide limited information about modelled loudness or acoustic context. Complementary approaches are often analysed separately, making their temporal relationships difficult to examine. This conceptual paper proposes a human-centred multilayer framework for continuous HAE characterisation. It organises signal-derived information into three layers: (1) a Physical Layer describing measured acoustic characteristics; (2) a Perceptual Layer representing the temporal evolution of loudness using the Moore-Glasberg-Schlittenlacher method; and (3) a Context Layer assigning each interval to the broad acoustic category it most closely resembles. Timestamps align the outputs while preserving their definitions, temporal resolutions and uncertainties. Local processing retains descriptors and assignments without storing raw audio or extracting linguistic content. Intended users, clinical activities and care objectives guide monitoring design and interpretation, while associations with human responses or clinical outcomes require independent evidence. The framework does not combine the layers into a single score or treat them as direct measures of experience or clinical effect. It provides a traceable basis for examining complementary acoustic evidence over time. Technical and ecological validation are required to establish its accuracy, generalisability and practical value.

Corresponding author: *a.solorioresendiz1@universityofgalway.ie*

Keywords: *healthcare acoustic environments, continuous acoustic monitoring, modelled loudness, audio classification, human-centred design*

1. Introduction

Healthcare acoustic environments (HAEs) are shaped by sounds from clinical procedures, alarms, speech, equipment, movement and building services. These sounds form part of routine care: some support communication, situational awareness and clinical action, whereas others may disturb patient rest or increase demands on staff [1], [2], [3]. Their meaning and effects depend on when they occur, who is exposed to them, the activity taking place and the clinical purpose they serve. Consequently, HAE quality cannot be judged only by how quiet an environment is; meaningful characterisation must consider its acoustic characteristics together with the care context.

Despite this complexity, many HAE assessments rely primarily on sound-level descriptors such as sound pressure level (SPL), A-weighted equivalent continuous sound level (LAeq), maximum level and percentile levels, often supplemented by temporal and spectral descriptors. These measures quantify acoustic magnitude and variation and support reproducible comparisons across rooms, units, periods and interventions [4], [5]. They can show whether predefined criteria are exceeded, but compliance with a numerical criterion does not establish whether the acoustic characteristics are appropriate for a particular listener, activity or care objective [6], [7].

Soundscape research provides a conceptual basis for this distinction by relating an acoustic environment to how it is perceived, experienced or understood by people in context [8]. This perspective does not make physical descriptors less important. It shows why sound levels alone cannot characterise perceived quality or contextual appropriateness.

Healthcare research has extended HAE assessment in several directions. Source-specific psychoacoustic analyses have examined modelled annoyance and intrusiveness [9], while clustering methods have been used to distinguish source related patterns [10]. Conceptual work has linked sound level measurement with experience and interpretation [11], and machine listening methods have been applied to neonatal sound sources and overlapping medical alarms [12], [13]. Taken together, these approaches provide complementary evidence, but the temporal relationships between their outputs remain difficult to examine.

To address this limitation, this conceptual paper proposes a human-centred multilayer framework for continuous HAE characterisation. It organises complementary signal-derived evidence within a common temporal representation. Drawing on relevant principles of ISO 9241-210 [14], the framework is human-centred in the specific sense that intended users, their tasks and goals, the care environment and the monitoring objective guide its configuration and interpretation. It remains signal-based and soundscape-informed and does not directly measure patient or staff experience. Relationships with human responses, clinical outcomes or operational data require external evidence. Technical and ecological validation are needed to establish accuracy, generalisability and practical value.

2. Multilayer framework for HAE characterisation

2.1 Framework overview

The framework distinguishes signal-derived representations from their interpretation within a healthcare setting. This distinction is consistent with conceptual work showing that physical levels alone do not capture how an acoustic environment is perceived or understood in context [15], [16]. As shown in Figure 1, its analytical core comprises three complementary layers:

- The Physical Layer describes measured acoustic characteristics at the microphone position, including background conditions.
- The Perceptual Layer calculates modelled loudness over time using the Moore-Glasberg-Schlittenlacher method specified in ISO 532-3 [17].
- The Context Layer assigns each analysed interval to a broad acoustic category and reports the associated classification score.

The Context Layer should not be confused with the healthcare context, which remains external to the analytical core. The healthcare context includes the clinical space, intended users, tasks and activities, care objectives and observation period. It defines the monitoring objective and guides measurement position, descriptor and category selection, aggregation periods and the external evidence needed for interpretation and validation. These elements must be documented independently, not inferred from the acoustic signal. The framework therefore assigns common analytical roles without prescribing universal descriptors or categories.

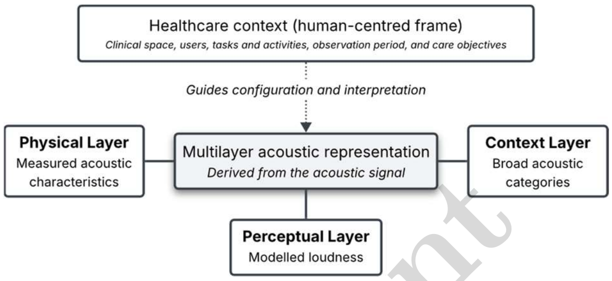

Figure 1. Conceptual structure of the proposed framework.

Although the layers serve distinct analytical functions, their outputs are not necessarily statistically independent because they derive from the same acoustic signal and may use related temporal and spectral information. Each retains its appropriate temporal resolution; timestamps allow alignment without identical update rates.

Aligned outputs can be examined together but do not establish the clinical purpose, appropriateness or effects of the represented sounds. Alignment creates no hierarchy, causal pathway or composite score.

2.2 Physical Layer: measured acoustic characteristics

The Physical Layer provides a quantitative representation of the acoustic characteristics measured at the microphone position. Its descriptors are calculated from a calibrated acoustic signal and may characterise sound level, temporal variation, background characteristics, time-weighted maxima and spectral content. Depending on the monitoring objective, these may include L_{Aeq}, L_{A10}, L_{AFmax}, L_{ASmax}, temporal level profiles, background estimates, peak indicators and spectral descriptors. The framework does not prescribe a fixed set; descriptor selection should reflect the clinical setting, relevant timescales and intended comparisons.

This multidimensional representation is useful because acoustic characteristics in healthcare spaces vary over time [4], [5]. A single L_{Aeq} calculated over a long period may conceal differences between sustained background sound, brief peaks, recurring patterns and periods of increased activity. Two observation periods may therefore have similar L_{Aeq} values while differing in background variability or spectral content. Retaining these characteristics separately provides a more informative account of how the measured acoustic environment changes over time.

Interpretation depends on how and where the measurements were obtained. Calibration, microphone position and orientation, acquisition settings, temporal resolution, aggregation periods and equipment performance should therefore be documented. A single microphone represents acoustic characteristics at one position and cannot be assumed to describe an entire room or the exposure of every occupant. Claims about individual exposure require a defined receptor position and observation period, together with evidence that the measurement adequately represents the acoustic characteristics encountered by that person.

Within these limits, the Physical Layer supports comparisons between locations and observation periods, identification of temporal trends, analysis of background characteristics, and evaluation of changes following acoustic or operational interventions. Source attribution, estimation of modelled loudness and assessment of effects on patients, staff or visitors require evidence beyond this layer.

2.3 Perceptual Layer: modelled loudness

The Perceptual Layer is deliberately restricted to modelled loudness. Its principal outputs are short-term loudness (STL) and long-term loudness (LTL), calculated using the Moore-Glasberg-Schlittenlacher method specified in ISO 532-3 [17]. In this framework, *perceptual* refers specifically to modelled loudness and does not encompass the broader experience or appraisal of the HAE.

STL represents the momentary loudness of a short sound segment and responds relatively quickly to temporal variation. LTL is derived by further smoothing STL and represents a longer segment, so it changes more slowly. Both vary over time; LTL is not equivalent to a shift- or day-level average. Sounds with similar L_{Aeq} values may differ in spectral content and temporal structure, producing different loudness patterns. These outputs complement the energy-based descriptors provided by the Physical Layer.

The outputs depend on the acoustic input, calibration, frequency response and listening configuration used in the calculation. They estimate loudness for an adult listener with no hearing problems under specific listening conditions. Individual hearing characteristics, health status, attention, activity and clinical circumstances fall outside this representation.

For continuous monitoring, STL and LTL time series may be aligned with the other layers. ISO 532-3 does not specify a single value for the overall loudness of a time-varying signal lasting more than 5 s. Summaries over longer observation periods are therefore post-processing choices, not direct ISO outputs. The aggregation method and loudness quantity must be stated explicitly, and the temporal trace retained when short-lived changes are relevant.

This layer does not directly measure annoyance, comfort, stress, disturbance, soundscape quality or clinical effect. Relationships with these constructs require independent human or clinical evidence.

2.4 Context Layer: classification of broad acoustic categories

The Context Layer produces a time sequence of broad acoustic category assignments. In the present configuration, the model evaluates each interval against four categories: speech, alarms, impulsive sounds, and mechanical /equipment sounds. It retains the highest-scoring category and its associated classification score. The classifier is non-polyphonic: an assignment indicates which category the interval most closely resembles, not every sound that is present. It does not identify a discrete event or estimate its onset and offset. The category set is not universal and may be modified for a particular healthcare setting or monitoring objective, provided that the corresponding model is developed and validated.

Alignment with the other layers allows measured acoustic characteristics and modelled loudness to be examined in relation to the assigned categories. For example, a change in level or loudness during an interval assigned to the alarm category can be identified for further interpretation, but the assignment does not establish that an alarm caused the change. Suitable analyses include the proportion and temporal distribution of category assignments, consecutive assignments, transitions between categories and the acoustic characteristics associated with each category. Consecutive assignments should not be treated as counts or durations of detected events.

Healthcare machine-listening studies have addressed source classification and overlapping medical alarms [12], [13]. Performance may nevertheless vary with room acoustics, equipment, overlapping sounds and differences between the training and deployment data. Category definitions, model and software versions, the classification rule, score interpretation and validation results should therefore remain traceable. A classification score should not be interpreted as a calibrated probability unless calibration has been demonstrated. Each assignment is a model inference, not an observed clinical fact.

An assignment to the speech category means that this category received the highest score for the interval. It does not involve speech recognition, transcription, speaker identification or linguistic analysis. Similarly, assignment to an alarm or equipment category does not determine the identity, urgency, function or clinical necessity of the corresponding sound. These interpretations require independent evidence, such as observations, alarm logs or workflow information.

3. Framework implementation and traceability

A practical implementation processes the calibrated acoustic signal locally using methods appropriate to each layer. In the present implementation, physical descriptors and acoustic category assignments are generated at one-second intervals. Loudness is calculated at the temporal resolution required by ISO 532-3:2023 [17] before any documented summarisation. Timestamps align the outputs without requiring identical update rates. Figure 2 summarises this workflow.

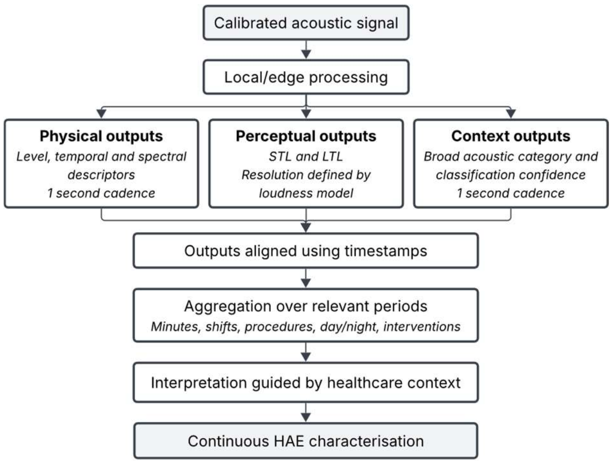

Figure 2. Local processing workflow and timestamp alignment of the three analytical layers before context-guided aggregation.

Aligned outputs may be summarised over periods defined by the monitoring objective, such as minutes, shifts, procedural phases, interventions or day and night periods. Sound levels require energy-based aggregation. Loudness measures covering longer periods are documented post-processing quantities because ISO 532-3:2023 [17] does not define a single overall loudness for time-varying signals longer than 5 s. Category assignments may be described by prevalence, temporal distribution and transitions, but successive labels must not be interpreted as event counts or durations. For each interval, only the highest-scoring category and its classification score are retained.

Table 1 summarises the outputs and possible analytical uses of the three layers. Timestamp alignment allows physical descriptors and modelled loudness to be examined for intervals assigned to a given acoustic category. Independently documented clinical information can then define comparison periods and guide interpretation.

Continuous monitoring of the HAE also requires data minimisation. Local processing can generate the proposed outputs without retaining raw audio. Stored data may be restricted to physical descriptors, loudness values, category assignments, classification scores and temporal summaries. A speech assignment contains no transcription, speaker identity or semantic information. Nevertheless, timestamped patterns may reveal clinical activity, so access, retention, security and appropriate use must be addressed in the monitoring design.

Table 1. Outputs and possible analytical uses of the three layers.

Layer	Output	Potential analytical use
Physical	Calibrated level, temporal, background, peak and spectral descriptors at the microphone position	Compare acoustic characteristics across locations and periods; identify temporal trends; evaluate changes during interventions
Perceptual	STL and LTL values, with explicitly defined loudness summaries	Examine brief and sustained changes in modelled loudness; compare patterns across activities, periods and settings
Context	Broad acoustic category assigned to each interval and associated classification score	Describe the distribution of category assignments and transitions between them; relate assigned categories to physical and loudness conditions

Reports should document measurement position, calibration, acquisition settings, temporal resolutions, aggregation periods, preprocessing, treatment of missing data and time synchronisation. Context Layer reporting should additionally state the category definitions, model and software versions, decision rule, interpretation of the classification score and relevant validation results. These details support traceable comparisons across rooms, units, studies and interventions.

Applicable standards can guide measurement and reporting but citing them does not establish conformity. IEC 61672-1:2013 [18] specifies performance requirements for sound level meters and can inform instrument selection and system evaluation. IEC 60942:2017 [19] specifies performance requirements for sound calibrators. ISO 1996-1:2016 [20] defines basic quantities and assessment procedures for environmental noise and does not prescribe noise limits. ISO 532-3:2023 [17] specifies the Moore-Glasberg-Schlittenlacher method for estimating the loudness and loudness level of stationary and time-varying sounds. None of these standards defines the proposed acoustic categories, which therefore require independent operational definitions and validation.

4. Implications, limitations and future validation

Research on HAEs includes sound-level monitoring, assessments of perceived quality and automated approaches to acoustic content [5], [6], [10], [12]. These approaches yield different forms of evidence and should not be treated as interchangeable. The framework's contribution is organisational: it places complementary signal-derived outputs on a common timeline, preserving their distinct meanings and allowing independently collected human and clinical observations to guide interpretation.

This separation is central to the framework's human-centred orientation. Here, human-centred refers to defining monitoring objectives and interpreting outputs in relation to users, clinical activities and care objectives; it does not add another measurement layer. Although developed for interactive systems, ISO 9241-210:2019 [14] offers relevant principles for grounding these decisions in the intended users and context of use. The practical relevance of an acoustic pattern therefore depends on its relation to the activities and needs of the people using the space [16].

At this stage, the aligned outputs should not be interpreted as direct measures of individual acoustic exposure or subjective response, nor as detections of discrete sound events. Their interpretation and generalisability remain conditional on measurement position, model assumptions and classifier performance in the deployment environment. These boundaries define the technical and ecological validation required before clinical relevance can be assessed.

The next phase, planned with clinical partners at University Hospital Galway, Ireland, will characterise the technical performance of the framework before ecological relationships are examined. Physical outputs will be compared with suitable reference measurements, and loudness calculations will be checked against an independent implementation of ISO 532-3:2023 [17] using defined test signals. Context assignments will be compared with reference annotations, with results reported by category and setting. The relationship between classification score and observed correctness, the effect of overlapping sounds and performance differences between development and deployment data will be assessed separately.

Ecological validation will then examine how the aligned outputs relate to independently collected human, clinical or operational observations. Subjective data collection may draw on ISO 12913-2:2018 [21] when adapted to healthcare, and relevant outcomes and observation periods should be specified in advance according to the monitoring objective. The central test is whether the combined representation improves the interpretation of these observations or the evaluation of interventions compared with individual descriptors. This responds to calls for stronger links between measured acoustic characteristics and health-related outcomes and for longer-term intervention evaluation [3].

Implementation should begin with a defined monitoring objective from which measurement and analysis choices are derived. Because the acoustic categories are model outputs, changing their set requires corresponding model development and validation. This objective-led approach translates the framework's human-centred orientation into a testable monitoring design.

5. Conclusion

This paper proposes a human-centred multilayer framework for the continuous characterisation of HAEs. The framework brings physical descriptors, modelled loudness and broad acoustic category assignments into a temporally aligned representation while retaining the definition, resolution and uncertainty of each output. This provides a basis for examining acoustic characteristics in relation to clinical activities, user needs and care objectives without treating signal-derived outputs as proxies for individual exposure, subjective response or clinical effect. The proposal now requires layer-specific technical validation and ecological validation in healthcare settings. Its practical value will depend on whether the combined representation supports more informative interpretation and evaluation of HAEs than individual descriptors used alone.

References

[1] K. Hummel, B. R. Lowndes, E. Ryherd, and V. Kennel, "The Impact of Soundscapes on Healthcare Teams: A Literature Review," in *Proceedings of the Human Factors and Ergonomics Society*, SAGE Publications Inc., 2024, pp. 1777–1782. doi: 10.1177/10711813241269257.

[2] W. Wang, R. Han, J. Li, Y. Wang, and Y. Zhang, "Workload-moderated pathways from hospital acoustic environment to healthcare workers' health and burnout via noise stress," *Journal of Building Engineering*, vol. 123, Apr. 2026, doi: 10.1016/j.jobe.2026.115931.

[3] E. Han, H. Kang, and Y. Jang, "Noise reduction interventions in intensive care units: a systematic review," Feb. 01, 2026, *Churchill Livingstone*. doi: 10.1016/j.iccn.2025.104234.

[4] C. Jerlehag, P. J. Lee, S. H. Park, T. Jones, and N. Carroll, "Acoustic environments of patient room in a typical geriatric ward," *Applied Acoustics*, vol. 133, pp. 186–193,

Apr. 2018, doi: 10.1016/j.apacoust.2017.12.022.
[5] J. L. Darbyshire, P. R. Greig, L. Hinton, and J. D. Young, "Monitoring sound levels in the intensive care unit: A mixed-methods system development project to optimize design features for a new electronic interface in the healthcare environment," *Int. J. Med. Inform.*, vol. 153, p. 104538, Sep. 2021, doi: 10.1016/j.ijmedinf.2021.104538.
[6] B. Lam *et al.*, "Assessing the perceived indoor acoustic environment quality across building occupants in a tertiary-care public hospital in Singapore," *Build. Environ.*, vol. 222, p. 109403, Aug. 2022, doi: 10.1016/j.buildenv.2022.109403.
[7] J. Yang, Y. Wu, and S. T. Ng, "From compliance to perception: A systematic review and quantitative synthesis of indoor environmental quality in healthcare facilities," *Results in Engineering*, vol. 29, Mar. 2026, doi: 10.1016/j.rineng.2026.109534.
[8] International Organization for Standardization, *ISO 12913-1:2014 - Acoustics - Soundscape - Part 1: Definition and conceptual framework*. Geneva, 2014.
[9] D. De Salvio, V. Amodeo, M. Garai, and S. Secchi, "Noise Assessment Within Healthcare Facilities: Annoyance And Intrusiveness Of Sound Sources," European Acoustics Association, Dec. 2025, pp. 3933–3938. doi: 10.61782/fa.2025.0607.
[10] M. Cingolani, D. De Salvio, D. D'Orazio, and M. Garai, "Clustering analysis of noise sources in healthcare facilities," *Applied Acoustics*, vol. 214, p. 109660, Nov. 2023, doi: 10.1016/j.apacoust.2023.109660.
[11] J. Mackrill, R. Cain, and P. Jennings, "Proposing a Conceptual Framework to Develop the Hospital Soundscape Through Visual Communication," *Design Journal*, vol. 19, no. 3, pp. 491–509, May 2016, doi: 10.1080/14606925.2016.1149330.
[12] M. Tailleur, V. Lostanlen, J.-P. Rivière, and P. Aumond, "Machine listening in a neonatal intensive care unit," in *Detection and Classification of Acoustic Scenes and Events 2024 (DCASE2024)*, Tokyo, 2024, pp. 166–170. [Online]. Available: http://arxiv.org/abs/2409.11439
[13] K. Kishimoto *et al.*, "Detection of Polyphonic Alarm Sounds From Medical Devices Using Frequency-Enhanced Deep Learning: Simulation Study," *JMIR Med. Inform.*, vol. 13, 2025, doi: 10.2196/35987.
[14] International Organization for Standardization, *ISO 9241-210 - Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems*. Geneva, 2019.
[15] F. Aletta, J. Kang, and Ö. Axelsson, "Soundscape descriptors and a conceptual framework for developing predictive soundscape models," May 01, 2016, *Elsevier B.V.* doi: 10.1016/j.landurbplan.2016.02.001.
[16] J. Chen and H. Ma, "A conceptual model of the healthy acoustic environment: Elements, framework, and definition," *Front. Psychol.*, vol. 11, p. 554285, Oct. 2020, doi: 10.3389/fpsyg.2020.554285.
[17] International Organization for Standardization, *ISO 532-3 - Acoustics - Methods for calculating loudness - Part 3: Moore-Glasberg-Schlittenlacher method*. Geneva, 2023.
[18] International Electrotechnical Commission, *IEC 61672-1:2013 - Electroacoustics - Sound level meters - Part 1: Specifications*. Geneva, 2013.
[19] International Electrotechnical Commission, *IEC 60942:2017 - Electroacoustics - Sound calibrators*. Geneva, 2017.
[20] International Organization for Standardization, *ISO 1996-1 - Acoustics - Description, measurement and assessment of environmental noise - Part 1: Basic quantities and assessment procedures*. Geneva, 2016.
[21] International Organization for Standardization, *ISO 12913-2 - Acoustics - Soundscape- Part 2: Data collection and reporting requirements*. Geneva, 2018.