



WHITEPAPER:

Preserving Privacy in Clinical Trials: Scientific, Ethical, and Operational Considerations of On-Device Cough Monitoring

The Hyfe Team

April 2025

hyfe.com

Executive Summary

The integration of digital technologies into clinical trials has ushered in a new era of continuous, high-resolution health monitoring, offering unprecedented scientific and operational advantages over traditional and incumbent methods. However, the extensive use of digital tools – in particular audio recording technologies – raises critical concerns around participant privacy, data security, trust, and the accuracy and validity of clinical trial outcomes.

This paper provides a thorough exploration of the imperative to preserve participant privacy within the context of clinical trials, particularly emphasizing the unique challenges posed by audio-based monitoring tools. The analysis underscores that maintaining robust privacy standards is not merely an ethical and regulatory requirement but a *crucial factor in the scientific integrity and operational success of modern clinical research*.

Privacy concerns significantly impact participant trust and trial retention, exacerbating dropout rates and potentially biasing trial outcomes. Regulatory frameworks – such as GDPR and HIPAA – categorically recognize audio and biometric data as highly sensitive, demanding rigorous privacy safeguards and explicit consent processes. Continuous monitoring methods that rely on traditional audio recording pose substantial privacy risks, create operational burdens due to high data volumes, and can inadvertently influence participant behavior (the Hawthorne effect/ white coat effect), thereby jeopardizing trial validity.

The paper highlights solutions through the adoption of edge computing and privacy-preserving technologies. Specifically, on-device audio processing technologies, such as those implemented in a new generation of platforms that include Hyfe's CoughMonitor Suite (CMS) or Actigraph's LEAP devices, enable accurate and reliable health monitoring without recording or transmitting identifiable audio data.

These technologies leverage advanced AI algorithms that process audio data locally, instantly discarding all audio after analysis. This approach dramatically reduces privacy risks, enhances regulatory compliance, and significantly lowers logistical and operational burdens associated with traditional monitoring techniques.

A detailed case study of Hyfe's CMS demonstrates the viability and efficacy of privacy-centric, on-device cough monitoring in clinical trials. This approach has been validated through rigorous scientific testing, demonstrating accuracy comparable to traditional methods but with enhanced participant trust and improved retention rates.

This paper argues that prioritizing privacy through technologically sophisticated, privacy-preserving digital solutions is not only ethically essential but also scientifically advantageous. As clinical trials increasingly rely on digital endpoints and decentralized models, technologies built with privacy at their core will inevitably dominate the clinical research landscape, aligning rigorous scientific standards with ethical responsibilities to foster participant trust, operational effectiveness, and data integrity.

Table of Contents

| | |
|--|----|
| Executive Summary | 2 |
| Table of Contents | 4 |
| Introduction | 6 |
| Overview of Digital Tools Used in Clinical Trials | 7 |
| 1. Wearable and Biosensor-Based Tools | 7 |
| 2. Smartphone-Based Tools | 8 |
| 3. Environmental and Digital Biomarker Tools | 8 |
| 4. AI and Analytics Platforms | 8 |
| Privacy in Clinical Trials: Principles and Regulations | 10 |
| Ethical Foundations | 10 |
| Regulatory Frameworks | 11 |
| GDPR, HIPAA, and global equivalents | 11 |
| Specific requirements for audio and voice data | 12 |
| Practical considerations | 12 |
| Scalability and Ease of Use | 12 |
| Cost | 13 |
| Low Friction | 13 |
| Relevance to Continuous Monitoring | 13 |
| Guidance from Regulatory Bodies | 14 |
| An overview of Current Practices in Clinical Trials Monitoring | 14 |
| Technical Implications of Audio Recording in Clinical Trials | 16 |
| The Hawthorne Effect and Behavior Modification | 16 |
| Evidence of Altered Behavior | 17 |
| Impact on Endpoint Validity and Generalizability | 17 |
| High Attrition Rates due to Participant Discomfort | 18 |
| Data Minimization and Risk Surfaces | 18 |
| Risks of Recording and Storing Raw Audio | 18 |
| Attack Surfaces for Re-identification and Unintended Uses | 18 |
| Logistical Burdens | 19 |
| Storage, Bandwidth, and Compute Requirements | 19 |
| Device Design Limitations | 20 |
| Infrastructure Requirements and Complexity at Scale | 20 |
| Edge Computing and On-Device Analytics: A New Paradigm | 20 |
| What is Edge Computing in Clinical Monitoring? | 21 |
| Definition and Technical Principles | 21 |
| Scientific Validity of On-Device Signal Processing | 22 |
| Accuracy and Reliability Without Audio Storage | 22 |

| | |
|--|-----------|
| Benefits Beyond Privacy | 23 |
| Extended Battery Life/ Low compute requirements | 23 |
| Lightweight and Ergonomic Hardware | 23 |
| Real-Time Processing and Feedback | 24 |
| Low Friction, Enhanced User Experience | 24 |
| Case Study: Cough Monitoring Without Audio Recording | 24 |
| Use Case: Continuous Cough Monitoring in Respiratory Trials | 25 |
| The Need for Frequent, Unobtrusive Monitoring | 25 |
| Traditional Methods and Their Limitations | 25 |
| Hyfe's On-Device Cough Detection Approach | 27 |
| Overview: A Privacy-First, Scalable Solution | 27 |
| Algorithmic Architecture | 27 |
| Validation and Performance Metrics | 28 |
| Privacy Design Choices and Their Tradeoffs | 29 |
| Core Privacy Features | 29 |
| Benefits | 29 |
| Tradeoffs | 30 |
| The Ethical Advantage of Designing for Privacy | 30 |
| Rebuilding Participant Trust | 30 |
| Enhancing Trial Recruitment and Retention | 31 |
| Implications for Broader Clinical Trial Design | 32 |
| Generalizability of the Approach | 32 |
| Design for Scale: Larger Trials, Longer Duration | 33 |
| Future-Proofing Trial Infrastructure | 33 |
| Limitations and Tradeoffs | 34 |
| Edge AI Constraints | 34 |
| Update Mechanisms and Model Improvements | 34 |
| Edge Device Limitations in Handling Complex Contextual Logic | 35 |
| Clinical Trial-Specific Constraints | 35 |
| No Ground Truth Data Retained | 35 |
| Tradeoffs in Scientific Interpretation and Endpoint Design | 36 |
| Balancing Privacy, Performance, and Complexity | 36 |
| References | 37 |

Introduction

Clinical research is experiencing a dramatic shift toward integrating digital tools that enable continuous, objective, and high-resolution monitoring of health indicators¹. Wearable devices, smartphone sensors, and audio recording systems now regularly collect sensitive personal data, promising unprecedented scientific insights and improved clinical outcomes². Yet, these powerful tools carry significant ethical, regulatory, and practical challenges³. A top priority is the urgent need to balance comprehensive data collection with strong safeguards for participant privacy.

The collection of audio data, including speech and cough or other respiratory sounds, represents a particularly sensitive area due to its inherent identifiability and intimate nature. Participants often perceive audio recording as intrusive, heightening concerns about surveillance, misuse, and behavioral alterations such as the Hawthorne effect⁴. These concerns have not only ethical implications but also profound practical repercussions—participant distrust can increase dropout rates and introduce bias, undermining the very validity of clinical trial results⁵.

In response to these challenges, this whitepaper explores a new paradigm for clinical monitoring: *privacy-preserving, on-device audio processing and cough monitoring*. Focusing specifically on cough monitoring as a critical case study, we examine how edge computing and advanced AI techniques enable *rigorous and precise health assessments without compromising participant privacy*. By processing data locally on participant devices and discarding raw audio

¹ Topol, 2019; Harrer et al, 2019

² Tackney, Steele et al. (2024), Coravos, Goldsack et al. (2019), Nebeker et al. (2019)

³ Narayanasetty, Swaroop & Jallu, Dr. (2021)

⁴ Nebeker et al, 2017; Kostick et al, 2019; OHRP, 2018

⁵ Ostherr, K. et al. (2017)

immediately, these technologies represent a significant advancement in privacy-by-design methodologies.

By exploring ethical principles, regulatory frameworks (such as GDPR and HIPAA), and the practical aspects of scaling clinical trials, we offer a comprehensive guide to conducting privacy-centric clinical research.

Ultimately, by highlighting Hyfe's innovative approach to on-device cough monitoring, this whitepaper seeks to demonstrate that scientific rigor, ethical responsibility, and participant trust are not mutually exclusive – in fact, they can and must coexist for the future of digital clinical trials to realize their full potential.

Overview of Digital Tools Used in Clinical Trials

Digital technologies are transforming the design, execution, and evaluation of clinical trials. These tools enhance data resolution, improve patient adherence, and enable decentralized study (DCT) models⁶. Here is a high-level overview of key categories in current use.

1. Wearable and Biosensor-Based Tools

Wearables and biosensors are used to collect physiological and behavioral data passively and continuously. These include smartwatches, chest straps, adhesive patches, and implantables. Common endpoints measured include heart rate, physical activity, sleep, oxygen saturation, and respiratory rate. Their utility spans chronic disease management, evaluating cardiac safety, and remote patient monitoring (RPM).

⁶ Fagherazzi, et al. (2024) argue that integrating voice monitoring in trials could improve patient monitoring and decentralization of trials. However, it stresses “privacy and security by design” as paramount to earning patient buy-in and recommends robust privacy safeguards (local processing, secure storage) to ensure participants and clinicians accept voice-based tools.

2. Smartphone-Based Tools

Smartphones are used to capture self-reported outcomes, monitor adherence to drug intake, deliver digital interventions, and collect passive data such as GPS, screen use, and app interactions. They also support real-time assessments through ecological momentary assessment (EMA) tools. Smartphones are especially valuable in delivering digital therapeutics (DTx), mental health interventions and research, and for hybrid trial designs.

3. Environmental and Digital Biomarker Tools

These tools measure external exposures and ambient factors, or derive biomarkers from digital traces. Examples include air quality monitors, ambient noise sensors, and devices that assess environmental triggers of disease. Digital environmental biomarkers – features derived from digital interactions or physiological signals – offer novel endpoints in areas like neurology, pulmonology, and behavioral sciences.

4. AI and Analytics Platforms

Artificial intelligence (AI) systems process large volumes of clinical trial data for signal detection, predictive modeling, and endpoint refinement. Machine learning (ML) is used to identify patterns, reduce noise, and generate insights from multimodal datasets. These platforms are central to risk-based monitoring, adaptive trial designs, and digital phenotyping.

| Category | Subcategory | Example Measurements |
|---|-----------------------|--|
| Wearable & Biosensor-Based Tools | Consumer Wearables | HRV, step count, SpO ₂ , sleep, stress, cough |
| | Medical-Grade Devices | ECG, respiration, temperature, hydration |

| | | |
|--|----------------------------------|---|
| | Patch-Based Sensors | Muscle activity, tremor, movement disorders Glucose Monitoring Devices – blood glucose, insulin levels |
| | Implantable/Ingestible Sensors | Medication adherence, pH, GI motility |
| | Smart Clothing | Posture, balance, fall risk detection |
| Smartphone-Based Tools | Mobile Health Apps | Patient reported outcomes (PROs), symptom logs, adherence tracking |
| | Sensor-Based Features | Gait, tremor, voice biomarkers, speech, cough |
| | Camera-Based Gait Analysis | Walking speed, stride length, fall detection |
| | Gamified Cognitive Tests | Attention, memory, reaction time |
| Environmental & Digital Biomarker Tools | Acoustic Monitoring | Cough, respiratory sounds, speech patterns and vocabulary |
| | Sleep & Activity Sensors | Sleep apnea, REM cycles, nocturnal movement |
| | Smart Home Devices | Daily routines, sleep/wake cycles, loneliness |
| | Environmental Exposure Trackers | Pollution, allergens, noise exposure |
| AI & Analytics Platforms | AI-Driven Remote Monitoring | Deterioration risk, treatment response prediction |
| | Digital Twin Models | Simulated control arms, trial forecasting |
| | Multimodal Integration Platforms | Cross-device signal fusion, real-time dashboards |

Fig 1 – High Level Overview of Digital Tools Used in Clinical Trials

Privacy in Clinical Trials: Principles and Regulations

Clinical research depends on the voluntary participation of individuals who consent to have their health data collected, analyzed, and sometimes shared. This transaction – between individual privacy and collective scientific benefit – is governed by both ethical principles and regulatory frameworks.

As technology enables more detailed and continuous forms of data collection, especially involving audio and voice, protecting participant privacy becomes a central challenge.

Ethical Foundations

- i. Autonomy and informed consent. Respect for autonomy is foundational in biomedical ethics⁷. In clinical trials, this principle is expressed through informed consent, which requires that participants are adequately informed of the nature, purpose, risks, and potential benefits of the study. Participants must understand what data will be collected, how it will be processed and stored, and for what purposes it may be shared.
- ii. Trust and voluntariness in participation. Clinical research depends on public trust. Voluntariness – free from coercion or undue influence – is central to ethical participation. Privacy violations, or even perceptions thereof, can erode trust and compromise both recruitment and retention. A study by Nass et al from 2009⁸ showed that confidentiality concerns were a leading reason for participant refusal in trials involving sensitive data. These

⁷ Nebeker, C. et al. (2017) highlight new privacy and ethics challenges posed by mobile apps, wearables, and sensing in research. Emphasizes need for updated consent practices and collaboration to protect participant privacy when using pervasive digital monitoring.

⁸ Nass SJ, Levit LA et al. (2009)

concerns become even more critical when using technology that records sound and / or voice of a trial participant.

- iii. Minimize harm. The principle of non-maleficence - “do no harm” - applies beyond physical risks. Psychological and social harms, including breaches of privacy, must be minimized. The Council for International Organizations of Medical Sciences (CIOMS) guidelines emphasize minimizing the impact of research procedures on participants, including data collection methods that may infringe on daily life or expose sensitive information⁹.

Regulatory Frameworks

GDPR, HIPAA, and global equivalents

The General Data Protection Regulation (GDPR) in the European Union¹⁰ and the Health Insurance Portability and Accountability Act (HIPAA) in the United States are two of the most influential data privacy frameworks also governing clinical trials. GDPR classifies health data and biometric data, including voice recordings, as “special categories” requiring explicit consent and strong protections¹¹. HIPAA, though less stringent, defines protected health information (PHI) broadly, covering any data that can identify an individual, including voiceprints¹².

Several other jurisdictions have adopted GDPR-like standards, including Brazil (LGPD), South Korea (PIPA), and parts of Canada (PIPEDA). These laws mandate transparency, purpose limitation, data minimization, and security safeguards.

⁹ Macrae, Duncan. (2007)

¹⁰ European Parliament and Council of the EU – GDPR (Regulation (EU) 2016/679), Recital 51, Art.9.

¹¹ Nautsch et al. 2019

¹² Isola et al (2025) confirms that under U.S. HIPAA, biometric identifiers like voiceprints are considered Protected Health Information (PHI). This means any audio recording that could identify an individual is subject to strict privacy safeguards and consent requirements. Highlights the importance of de-identification or avoidance of collecting such identifiers in research to stay HIPAA-compliant.

Specific requirements for audio and voice data

Audio (and voice) data, especially when processed with AI for features like cough detection or speech biomarkers, raise specific privacy concerns. It is not clear exactly what legal constraints apply to recording ambient sounds even if very short in duration. Such data are often identifiable, even in the absence of names or IDs, due to unique vocal characteristics. GDPR Recital 51 explicitly includes voice as biometric data requiring additional protection. HIPAA, while not biometric-specific, considers audio that can identify an individual as PHI¹³.

In practice, this means that continuous voice or cough monitoring technologies must implement on-device processing, avoid unnecessary storage, and anonymize or pseudonymize data when possible. Explicit consent is mandatory under GDPR, and in many cases, ethics boards will demand justification for such intrusive monitoring¹⁴.

Above mentioned realities mean that technologies relying on continuous audio recording for cough monitoring – such as some incumbent cough monitoring solutions that require human annotation or audio storage for later processing – will significantly limit the range of clinical trials where cough can be used as an endpoint, while also introducing additional challenges in study design and ethical approvals.

Practical considerations

Scalability and Ease of Use

Technological solutions that support privacy must also be scalable and user-friendly. High-friction consent processes or complex user interfaces can deter participation and increase dropout rates¹⁵. Privacy-preserving

¹³ The full text of Recital 51 can be found here: <https://gdpr-text.com/read/recital-51/>

¹⁴ European Commission: Ethics and Data Protection, 14 November 2018: https://commission.europa.eu/system/files/2020-06/5_h2020_ethics_and_data_protection_0.pdf

¹⁵ Turner, Birring (2023)

systems should be unobtrusive, ideally running in the background without requiring user interaction.

Cost

Implementing robust privacy safeguards can increase trial costs. Encryption, secure cloud infrastructure, legal compliance, and ethics consultations all add overhead. However, the cost of a privacy breach – financial, legal, and reputational – is typically far greater. Moreover, trust-based recruitment benefits from a strong privacy posture.

Low Friction

A growing area of innovation involves low-friction privacy techniques, such as on-device AI, federated learning, and edge computing¹⁶. These allow sensitive data (like audio from study participants) to be analyzed without leaving the participant's device, reducing exposure and enhancing compliance with privacy laws.

Relevance to Continuous Monitoring

Continuous cough monitoring, such as 24h sound recording for later human annotation, exacerbates these concerns. Even when data is anonymized, the feeling of being watched and recorded can reduce engagement or lead to biased behavior. Longitudinal trials must balance scientific needs with psychological comfort. Adaptive consent models, opt-in recording, and user transparency have been proposed as mitigations. The best solution seems to be on-device processing for cough timestamps, with no audio being recorded or stored.

¹⁶ De Brouwer, W. et al. (2021)

Guidance from Regulatory Bodies

The U.S. Department of Health and Human Services Office for Human Research Protections (OHRP) has cautioned that even when data is de-identified, persistent monitoring may alter behavior and affect voluntariness¹⁷. The Association of Clinical Research Organizations (ACRO)¹⁸ and various Institutional Review Boards (IRBs) have echoed the need for stronger justification and transparency when deploying surveillance-like technologies that record and store audio of trial participants.

An overview of Current Practices in Clinical Trials Monitoring

Traditional practices for monitoring symptoms like cough remain limited by outdated methodologies and privacy-related pitfalls. For instance, the traditional "gold standard" in cough monitoring relies on the use of continuous audio recording devices worn by participants, typically for 24-hour periods. These devices indiscriminately capture all surrounding audio, including conversations and ambient noise, which is then reviewed manually by human annotators to identify and count coughs.

These methods, while historically foundational for establishing objective cough metrics, present several critical limitations highlighted in many publications¹⁹. First, the privacy implications of recording all sound in a participant's environment are profound. Participants may feel surveilled or violated, particularly when recordings take place in private settings such as the home, during conversations or during sleep²⁰. Such

¹⁷ OHRP (2018)

¹⁸ ACRO White Paper (2020)

¹⁹ Turner, Birring (2023)

²⁰ Turner, Birring (2023) "The above-mentioned devices (...) are intrusive to the patient to some degree. The wearer is physically aware of being attached to the monitor, and whilst in use activities such as swimming and showering are restricted. The wearer may also be conscious of being recorded, potentially altering

discomfort has been associated with increased attrition rates and reduced engagement, ultimately threatening the integrity and statistical power of trials²¹. Additionally, devices that record and store audio carry external risks, such as potential access to audio files if the device is lost or stolen.

Second, the operational demands of these models are substantial. Manual annotation is labor-intensive, time-consuming, and costly. It is also inherently subjective, as inter-rater variability can affect consistency. The 24-hour constraint further limits the ecological validity of the data, failing to capture longer-term symptom patterns or variability across days and environments.²²

Other recent approaches have sought to mitigate these burdens by using devices that detect and transmit only "suspected sounds" to cloud servers for further analysis and sound classification. While this is a significant improvement on the continuous audio recording method, it does not fully address privacy concerns. Participants are still recorded without full control or knowledge of what is captured, and audio files must still be transmitted and stored, creating exposure risks and infrastructure dependencies. Moreover, this approach introduces technical and ethical complexities, including questions of data ownership, re-identification risks, and the opacity of cloud-based algorithms.

With both of these models, participants may alter their behavior due to the perceived presence of surveillance - suppressing coughs, avoiding speech, or changing environments - which introduces bias and compromises the generalizability of results.

the cough rate being measured—similar to the well-known phenomenon of 'white coat hypertension', whereby anxiety and other factors influence blood pressure when the subject is aware of being observed.

²¹ Henry SG et al. (2015)

²² Chung, Small et al. 2024

As such, incumbent solutions fall short of balancing the scientific demand for granular, longitudinal data with the ethical imperative to respect participant autonomy and privacy.

These limitations underscore the need for new models of digital clinical monitoring – that are passive, unobtrusive, scientifically rigorous, and fully privacy-preserving by design.

The emergence of on-device analytics and edge computing, already adopted on systems like Hyfe’s CoughMonitor Suite or [Actigraph’s LEAP devices](#), represents a promising step in this direction, offering an alternative that reduces privacy risk while enabling scalable and ethically aligned data collection.

Technical Implications of Audio Recording in Clinical Trials

Audio recording technologies offer new opportunities for clinical research. Voice, speech, and cough data can serve as biomarkers for neurological, cardiovascular, respiratory, and psychiatric conditions. However, embedding continuous or intermittent audio capture into clinical trials introduces significant technical and methodological challenges. These span from participant behavior modification to data security and infrastructure requirements. A rigorous understanding of these implications is essential for both trial design and ethical compliance.

The Hawthorne Effect and Behavior Modification

Evidence of Altered Behavior

The Hawthorne effect describes a participant's tendency to alter behavior due to awareness of being observed²³. This phenomenon has been documented across clinical and behavioral studies, particularly when monitoring tools are visible or personally intrusive²⁴²⁵²⁶²⁷. In trials involving audio capture – such as those evaluating novel antitussives like P2X3 inhibitors – participants may consciously or unconsciously alter their behavior, including how often they speak, suppressing cough symptoms, or changing social interactions. These behavioral shifts have been observed in both observational studies and user experience research, particularly in studies using home-based sensors or more intrusive audio-capturing technologies – such as chest-worn adhesive patches, belt-worn devices, or other setups that go beyond the simplicity of a wearable smartwatch.

Impact on Endpoint Validity and Generalizability

When participants behave differently under observation, clinical endpoints derived from audio – such as cough frequency, vocal biomarkers, or speech content, vocabulary used – may no longer reflect real-world patterns. This degrades internal validity and challenges generalizability. For example, if participants reduce coughing due to self-consciousness, an intervention may appear more effective than it actually is. In longitudinal trials, behavior modification may also wane over time, introducing time-based variability that is difficult to model or control for.

Mitigating the Hawthorne effect requires designing systems that are passive, discreet, and require no user interaction. Participants should not

²³ Berkhout et al. (2022)

²⁴ Adiyaman et al. (2015)

²⁵ Tanner et al. (2016)

²⁶ Nuredini et al, (2020)

²⁷ Barochiner et al. (2022)

feel surveilled. On-device processing and limited-data feedback loops can help technologies become more transparent and less intrusive over time.

High Attrition Rates due to Participant Discomfort

There is increasing evidence that perceptions of surveillance can lead to discomfort, distress, or withdrawal from clinical trials²⁸. A review by Nebeker et al. (2019)²⁹ found that digital health trials involving wearables or sensors had higher attrition rates when participants felt monitored or recorded continuously. In one qualitative study, participants described a “loss of personal space” when audio sensors were used in the home.

Data Minimization and Risk Surfaces

Risks of Recording and Storing Raw Audio

Raw audio data is both **rich in health-relevant features** and **inherently identifiable**. Voice patterns, ambient background sounds, and speech content can all contain sensitive information. Storing such data increases the risk of misuse – whether through accidental leakage, intentional breach, or secondary analysis not covered by original consent. Unlike structured health data, raw audio carries open-ended information that may reveal more than participants understand or anticipate.

Attack Surfaces for Re-identification and Unintended Uses

Even anonymized datasets containing voice data are susceptible to **re-identification**. Voice is a biometric identifier, and advances in

²⁸ Henry SG et al. (2015) – In a trial with optional audio-recording of doctor visits, only ~39% of visits were recorded (many declined due to confidentiality concerns). The low consent rate flags selection bias: certain patient groups (e.g. those with privacy worries) opt out, which can threaten validity.

²⁹ Nebeker et al. (2019)

machine learning (ML) make it possible to match de-identified recordings with public or previously collected voiceprints. Moreover, speech content may inadvertently include names, locations, or private health information. These vulnerabilities create a large “attack surface” for adversarial or negligent misuse.

In response, privacy-by-design principles recommend **data minimization** – collecting only the information necessary for a clearly defined clinical endpoint. Processing on-device and transmitting only derived metrics (e.g., cough counts, cough severity scores based on cough cluster or cough bouts dynamics) can dramatically reduce exposure. Where raw audio is essential, encryption at rest and in transit, role-based access control, and ethical firewalls against secondary use must be standard practice.

Logistical Burdens

Storage, Bandwidth, and Compute Requirements

Audio data is storage-intensive. One hour of uncompressed audio can exceed 600 MB; even compressed, continuous recording over weeks or months can generate terabytes of data per participant. This poses significant bandwidth and storage burdens, especially when trials involve hundreds or thousands of participants. Processing audio – whether for transcription, feature extraction, or machine learning purposes – requires additional compute resources, either on-device or in the cloud.

These demands translate into increased infrastructure costs, potential latency in analysis, and more complex regulatory compliance (e.g., under GDPR’s data localization requirements). Clinical trial sponsors must weigh whether the scientific value of raw audio justifies these costs, or whether summary features can suffice.

Device Design Limitations

Battery life, form factor, and compute capacity impose further limitations. Wearables capable of continuous audio monitoring must balance sensor fidelity with power consumption. Larger batteries may improve uptime but reduce comfort and wearability – potentially impacting adherence and data quality. Similarly, on-device processing reduces transmission risk but increases power draw, requiring more sophisticated engineering and thermal management.

Infrastructure Requirements and Complexity at Scale

At small scales, manual review, storage, and analysis of audio may be feasible. But as trials scale, exponential complexity emerges. Data pipelines must manage ingestion, quality control, encryption, metadata tagging, and compliant deletion. Cloud infrastructure must ensure uptime, redundancy, and jurisdictional compliance. Each layer of infrastructure introduces points of failure and additional obligations under data protection laws.

Moreover, scaling audio trials demands specialized skills – audio signal processing, edge computing, privacy engineering – that are not typical in most clinical research teams. Without careful planning, audio-enabled clinical trials may also fail due to logistical rather than scientific constraints.

Edge Computing and On-Device Analytics: A New Paradigm

As clinical trials increasingly integrate digital tools, a critical shift is occurring from centralized data capture toward decentralized, privacy-preserving clinical trials (DCTs) and analytics. One of the most promising approaches in this evolution is edge computing – a model in

which data is processed locally on a participant's device rather than being continuously recorded, transmitted, or stored in the cloud. For clinical trials involving acoustic signals such as cough, speech, or ambient sound, edge computing offers a solution that balances scientific rigor with participant autonomy, usability, and regulatory compliance. This also marks a significant departure from the historical "gold standard" in cough monitoring – continuous sound recording followed by manual cough counting through human analysis during later-stage processing.

What is Edge Computing in Clinical Monitoring?

Definition and Technical Principles

Edge computing refers to the execution of computational tasks – such as data filtering, feature extraction, and classification – directly on the device where data is generated³⁰. In contrast to traditional models, where data is recorded locally for later processing or transmitted to a centralized server or cloud for processing, edge systems analyze data at the source. In the context of clinical trials, this means that a wearable or smartphone can run an embedded analysis, for example, an AI model passively and continuously counting coughs, to detect events of interest without recording or storing any raw audio.

This represents a major departure from legacy methods, including:

- Audio recorders that capture continuous sound for later manual review (the historic standard for cough monitoring),
- Cloud-based systems that transmit and store large volumes of raw data for post hoc analysis, and
- Hybrid models that perform partial local processing but still require intermittent audio data uploads for full inference.

³⁰ Kotevska, Johnson et al. (2022)

In contrast, **fully edge-based systems** perform detection in real time, discard all audio, and generate only summary data – e.g., highly precise timestamps and event counts.

Scientific Validity of On-Device Signal Processing

Accuracy and Reliability Without Audio Storage

A primary concern in transitioning to on-device models is whether their outputs retain scientific validity. Specifically, can AI systems running on edge devices detect meaningful clinical events with sufficient accuracy, sensitivity, and specificity – without the need to store or inspect the underlying signal?

Recent work suggests the answer is yes. On-device acoustic models, such as those deployed in [Hyfe's Cough Monitoring Suite \(CMS\)](#), have demonstrated the ability to detect coughs in real time, in real world environments using low-power devices such as smartphones or wearable smartwatches. These models use a lightweight deep learning architecture optimized for low-latency inference, and operate in a fully privacy preserving way without storing or transmitting any raw audio.

A multicenter clinical validation study published in Nature's Scientific Reports³¹ evaluated the performance of a cough monitoring system deployed on a wearable smartwatch, and found high agreement between automatic on-device cough detection and expert-annotated audio recordings (sensitivity: 90%, 1 FP/hr). The automated system from Hyfe required no audio storage, relying instead on a real-time model trained to distinguish coughs from background noise and cough-like sounds.

Similar architectures have been validated outside of healthcare as well – for example, in wildlife monitoring and voice-controlled consumer

³¹ Chaccour et al, (2025)

electronics – where on-device inference has proven robust for keyword detection, speaker recognition, and other tasks in noisy environments.

These findings suggest that edge computing, when paired with task-specific acoustic AI, can meet the scientific and regulatory standards required for clinical trials.

Benefits Beyond Privacy

Extended Battery Life/ Low compute requirements

Processing data locally eliminates the need for constant wireless data transmission – a major source of energy drain in mobile and wearable devices. Edge models can run at low CPU usage and in intermittent bursts, preserving battery life. This enables long-term, uninterrupted monitoring, which is critical for trials measuring disease progression or treatment response over weeks or months.

Lightweight and Ergonomic Hardware

Because edge models reduce compute and storage burdens, they can be embedded in lightweight, ergonomic devices with minimal form factor. Participants are more likely to comply with monitoring protocols when devices are comfortable, silent, and easy to wear. Unlike bulkier systems requiring active user input or manual charging routines, edge-powered devices can be worn passively, even during sleep or physical activity.

Additionally, because of the low compute and memory requirements, powerful models can run on small, affordable devices, which facilitates scale for sponsors.

Real-Time Processing and Feedback

Edge systems enable immediate, local interpretation of data, opening new possibilities for real-time feedback and intervention. For example, in a behavioral cough suppression trial, a device could detect frequent coughing and prompt a therapeutic reminder without relying on cloud connectivity. This aligns with emerging models of digital therapeutics that aim to integrate sensing and intervention in a single, autonomous system.

Low Friction, Enhanced User Experience

The combination of privacy preservation, real-time operation, and device minimalism results in a low-friction experience for participants. There is no need to manage recordings, approve uploads, or troubleshoot connectivity. As a result, user burden is significantly reduced, which is correlated with higher adherence and lower attrition in digital trials.

Hyfe's experience with edge-based cough monitoring across diverse trial populations - ranging from patients with chronic respiratory conditions to those in post-viral recovery and in cardiovascular diseases like congestive heart failure - has consistently shown that minimal-intrusion systems are not only technically feasible but also more acceptable to participants and easier to scale for sponsors, in particular in large and decentralised clinical trials (DCTs).

Case Study: Cough Monitoring Without Audio Recording

New generation tools such as [Hyfe's CMS](#) exemplify a new generation of ethically aligned, technically sophisticated clinical monitoring platforms. By eliminating the need for audio recordings, these platforms meet the

growing demand for privacy, scalability, and continuous symptom tracking. Their success in both validation studies, industry-led adoption in drug development trials, and in real-world trials demonstrates that privacy and performance are not mutually exclusive – they can coexist through thoughtful design, robust modeling, and a clear understanding of clinical and participant needs.

As digital biomarkers, like digital cough monitoring, become integral to future trial protocols, systems like Hyfe’s CMS offer a blueprint for how to scale data collection without compromising ethics, comfort, or scientific rigor.

Use Case: Continuous Cough Monitoring in Respiratory Trials

The Need for Frequent, Unobtrusive Monitoring

As explored in previous sections, clinical trials in respiratory medicine increasingly demand objective, high-resolution symptom endpoints. To support these endpoints, continuous monitoring is often required – ideally over weeks or months.

While traditional cough monitoring solutions – such as audio recorders, manual annotations, or cloud-streamed microphones have allowed the field to progress significantly from Patient Reported Outcomes (PROs), they also introduce additional barriers to long-term adoption. These include privacy concerns, battery and storage limitations, high participant burden, new forms of bias in the output data as well as complex regulatory hurdles related to identifiable data.

Traditional Methods and Their Limitations

Historically, digital cough monitoring has relied on:

1. Manual diary-based systems supplemented by clinician interviews.
2. Tape recorders worn by participants, usually for a maximum of 24 hours of audio recording,
3. Cloud-connected microphones streaming raw audio for real-time classification.

These methods face multiple limitations, including but not limited to:

- a. Short recording windows (typically ≤ 24 hours) often fail to capture long-term variability and may miss to capture the mechanism of action of the tested intervention³²,
- b. Low ecological validity, as participants are likely to alter behavior while recorded,
- c. Manual review burdens, where recordings are laboriously annotated by human annotators with the potential risks of inconsistent inter-labeler concordance,
- d. Privacy objections, especially with in-home or overnight surveillance,
- e. Technical failures related to device fatigue, data loss, or participant dropout.

Anecdotal reports suggest participants tend to withdraw from clinical trials when they learn that their conversations, environmental sounds, and private moments may be recorded for future replay. This “surveillance effect” poses a major challenge to trial generalizability and data integrity.

³² Chung F, et al. (2024)

Hyfe's On-Device Cough Detection Approach

Overview: A Privacy-First, Scalable Solution

Hyfe's Cough Monitoring Suite (CMS) addresses the above mentioned limitations through a unique application of on-device acoustic AI. Rather than recording or streaming audio, Hyfe's CMS uses real-time edge processing to detect and count coughs locally on device, without ever recording, storing or transmitting any audio data. Only metadata – such as highly precise timestamped cough counts and device performance logs (e.g., battery life, wear/non-wear status) – is collected, minimizing both data exposure and operational complexity.

Hyfe's CMS is typically deployed via wearables (e.g., smartwatches) but can also be deployed via smartphones or any embedded devices, including custom integrations into third-party hardware via Hyfe's cough monitoring SDK and / or API.

The core value is continuous, passive, and privacy-preserving cough monitoring, scalable to thousands of participants across time zones and settings, easily performed over weeks, months or even years with minimal participant friction.

Algorithmic Architecture

Hyfe's detection engine is built on a lightweight convolutional neural network (CNN) based on Hyfe's proprietary dataset of billions of datapoints across multiple diverse environments, demographics, and languages. The model is optimized for:

- Low compute overhead, enabling real-time processing on mobile CPUs,
- Low power consumption, supporting 24/7 operation on wearables,
- On-device inference, meaning no audio ever leaves the device,

The processing pipeline includes:

1. Audio snippet sampling (~0.5–1s buffers),
2. Spectral transformation into Mel spectrograms,
3. Neural network inference, classifying each snippet as cough or not,
4. Post-processing filters to remove artifacts or duplicates,
5. Data logging, recording only the cough event and timestamp.

Hyfe's CMS model does not retain, buffer, or archive any audio snippets. It operates entirely in real time and is fully privacy preserving.

Validation and Performance Metrics

Hyfe's CMS has been independently validated through multiple studies, including:

- Chaccour C, et al. (2025) - Validates an on-device cough detector (Hyfe) against manual cough counts. In 546 hours of monitoring in real world environments as patients with cough due to many etiologies go about their daily activities, the wearable system detected coughs with ~90% sensitivity and a very low false-positive rate (~1 per hour), correlating almost perfectly ($r \approx 0.99$) with human annotators. Demonstrates that privacy-preserving, smartwatch-based cough monitoring can achieve clinical-grade accuracy. The study did not record or transmit identifiable audio on the Hyfe's system; cough events were detected locally.
- Galvosas M, et al. (2023) validated Hyfe's cough monitoring system deployed on a smartphone using solicited coughs in controlled environments. The study demonstrated a sensitivity of 91% and specificity of 98%, with a strong correlation between Hyfe's measured cough rate and that of human annotators (Pearson correlation: 0.968).

- Longitudinal cohort studies: Hyfe's cough monitoring tools have monitored thousands of participants for multiple weeks, across more than 50 academic and industry-sponsored trials, detecting billions of cough events. These studies demonstrated strong correlation between cough frequency trends and predictive patterns preceding disease onset or worsening and, most importantly, very high participant retention (>90%) in trials longer than 4 weeks³³.

These metrics confirm that scientifically valid insights can be derived without audio recordings, provided models are carefully trained, validated, and calibrated for the device and environment in use.

Privacy Design Choices and Their Tradeoffs

Core Privacy Features

Hyfe's CMS was designed around the principle of data minimization, as articulated in GDPR, HIPAA, and CIOMS guidelines. Key features include:

- On-device inference: No audio is recorded, saved, or transmitted.
- Ephemeral processing: Data is discarded after classification.
- Anonymized metadata: Logs contain only non-identifiable timestamps and counts.

Benefits

This architecture reduces:

- Regulatory burdens around identifiable or biometric data,
- Participant concerns about eavesdropping or surveillance,
- infrastructure costs related to cloud storage, encryption, and audits.

³³ For more details on many of these studies, see <https://www.hyfe.com/publications>

It also simplifies IRB approval processes, as many ethics boards consider audio recording a “higher-risk” activity requiring special justification.

Tradeoffs

However, this privacy-first approach introduces some limitations:

- No re-labeling: Events cannot be reviewed for manual quality control. Audio snippets cannot be used for further machine learning initiatives.
- Limited contextual data: Environmental cues and behavioural patterns are inaccessible.
- Model-dependent performance: All insights depend on real-time classifier accuracy.

To address these tradeoffs, Hyfe emphasizes robust model training, cross-validation across diverse populations, and transparent reporting of device performance. In some settings, hybrid models may still be required – for example, brief audio capture in pre-consented substudy arms for algorithm calibration and / or the use of a concurrent validation device in trials that require the retention of audio data.

The Ethical Advantage of Designing for Privacy

Rebuilding Participant Trust

Clinical research operates on a foundation of public trust – trust that data will be handled ethically, that personal boundaries will be respected, and that participation will not result in harm. In recent years, high-profile breaches, opaque consent processes, and intrusive digital

tools have strained that trust³⁴³⁵³⁶³⁷³⁸. Rebuilding it requires a structural shift: designing privacy into the system from the outset, not retrofitting it as an afterthought.

New generation approaches to privacy-by-design – such as those pioneered by Hyfe – reflect this shift. By creating tools that avoid audio recording and audio storage altogether, participants are offered a clear, easy-to-understand value proposition:

“your coughs will be counted, but your conversations, surroundings, and private moments will never be captured”.

This directly supports informed consent, which is only meaningful when participants truly understand what is – and is not – being done with their data.

Transparency and control are critical for participant empowerment. When users are offered the ability to verify how a system works, what data it collects, and how long that data persists, they are more likely to consent voluntarily and remain engaged over time.

Enhancing Trial Recruitment and Retention

There is growing evidence that privacy-centric design improves both recruitment and retention³⁹. Participants are more likely to enroll in studies that clearly explain data practices and limit personal exposure. This is particularly true in vulnerable populations – such as individuals with stigmatized conditions, adolescents, or those with prior negative healthcare experiences – where digital surveillance can produce a chilling effect, deterring engagement altogether.

³⁴ Yusuf, Dixon et al. (2024)

³⁵ van Rijssel, van Thiel et al. (2024)

³⁶ Ferretti & Vayena, (2022)

³⁷ Heidelberg, Kelman et al. (2020)

³⁸ Jumai Ehidiemen & Oladapo (2024)

³⁹ Henry et al. (2015)

Nebeker et al. (2019) found that participant attrition in mobile health studies was significantly higher when monitoring technologies were perceived as intrusive. Conversely, studies emphasizing autonomy and privacy-preserving features reported higher enrollment and longer engagement durations. In trials with low-friction, passive, and privacy-preserving designs, Hyfe has observed retention rates exceeding 90% over multi-week periods.

By aligning technological design with ethical norms, trials not only fulfill regulatory obligations – they also become easier to run, less likely to fail, and more reflective of real-world behavior.

Implications for Broader Clinical Trial Design

Generalizability of the Approach

While this case focuses on cough, the core principle – on-device analytics with no raw signal storage – is applicable across a wide range of biosignals. Heart sounds, respiratory sounds and respiratory rate, gait patterns, speech dynamics, and even facial micro-expressions can be detected using edge-based AI models. As sensor hardware becomes more capable and algorithms more efficient, a growing number of clinical endpoints will be measurable in a way that is both scientifically valid and ethically robust.

This model is particularly well-suited for decentralized and hybrid trials, where infrastructure varies widely, and where participant trust must be earned and sustained without in-person interaction.

Design for Scale: Larger Trials, Longer Duration

Edge-based systems also unlock new operational possibilities. By eliminating the need to transmit, store, or secure large volumes of raw data, the logistical burden of running long-term studies is dramatically reduced. Lower bandwidth and storage costs, reduced data auditing complexity, and decreased regulatory scrutiny all translate into greater scalability.

With less friction and more durable devices, participants can be monitored over longer periods, yielding more granular longitudinal data and allowing for the study of slow-moving conditions or rare symptom events. As trial cohorts grow in size and diversity, data from underrepresented populations can be more readily incorporated – enhancing external validity and equitable evidence generation.

Future-Proofing Trial Infrastructure

The clinical trial ecosystem is rapidly evolving, with growing expectations for real-world evidence, patient-centric design, and modular, interoperable systems. Privacy must be a foundational principle in this transformation.

Platforms such as Hyfe's CoughMonitor Suite demonstrate how future-ready tools can be both minimalist and powerful – capturing high-quality health and wellness signals with low participant burden, little infrastructure overhead, and high regulatory compatibility. Its modular architecture allows easy integration into other trial platforms, wearables, or therapeutic systems.

In this way, privacy-conscious, edge-based tools serve not only today's trials, but also anticipate the ethical and operational demands of tomorrow's research landscape.

Limitations and Tradeoffs

Edge AI is emerging as a transformative approach in clinical trials, enabling real-time, privacy-preserving analysis of biosignals directly on participant devices. Its benefits are significant: enhanced privacy, reduced infrastructure burdens, and improved participant compliance. However, edge computing is not without limitations. To fully realize its potential, trial designers must understand and plan for its constraints – especially when applied to complex, context-sensitive domains such as acoustic health monitoring.

Edge AI Constraints

Update Mechanisms and Model Improvements

One of the core limitations of edge-based AI systems is their limited capacity for dynamic model improvement. Unlike centralized models deployed on cloud servers, edge models do not automatically benefit from real-time updates, collective learning across users, or global error correction. Once deployed, the model that runs on a participant's device is static – unless explicitly updated via app or firmware versioning.

This introduces a challenge: how to continuously improve model performance across populations without storing sensitive data or requiring invasive update processes. While periodic model updates can be pushed as part of standard software maintenance, this approach lacks the immediacy of cloud-based retraining and requires careful version control across a distributed trial population.

Emerging techniques such as federated learning may mitigate this issue by enabling devices to contribute model updates without sharing raw data. However, these approaches are still under active research and come with their own complexity, especially in regulatory contexts.

Edge Device Limitations in Handling Complex Contextual Logic

Edge devices – particularly wearables and smartphones – are hardware-constrained. They are optimized for low power consumption, small memory footprints, and limited compute cycles. As a result, the models deployed to these devices must be compact and efficient, often at the cost of contextual complexity.

For example, a cough detection model operating on-device may accurately identify individual cough events but lack the context to determine clinical context: it may not infer whether the participant is in a conversation, sleeping, or exposed to an irritant – all of which could trigger and alter cough patterns. Without ambient or behavioral context, it becomes more difficult to distinguish signal from noise..

Clinical Trial–Specific Constraints

No Ground Truth Data Retained

Perhaps the most important methodological tradeoff in edge AI clinical trials is the absence of retained ground truth data. When models process acoustic signals in real time and discard raw audio, there is no way to retrospectively audit misclassifications or assess edge cases post hoc. This creates challenges for:

- Quality assurance and model debugging,
- Regulatory audits that request traceability of endpoint generation, and
- Scientific reproducibility, especially in high-stakes or pivotal trials.

However, this limitation can be strategically mitigated. One method is to deploy a ground truth validator – a parallel system that records brief audio segments under tightly controlled consent protocols. These validators can operate on a subset of participants or during a

predefined calibration period, offering a mechanism for model benchmarking and adjustment without undermining privacy at scale.

Another strategy is to pair on-device classifiers with synthetic test environments and high-quality reference datasets, ensuring robust validation prior to trial deployment. While this does not allow for participant-specific reanalysis, it supports regulatory defensibility and general performance assurance.

Tradeoffs in Scientific Interpretation and Endpoint Design

While edge AI can deliver high-frequency, low-friction symptom data, researchers must accept tradeoffs in interpretability. Without raw inputs, some types of secondary analysis – such as acoustic phenotyping, linguistic analysis, or source separation – are unavailable. This constrains the system to predefined endpoints (e.g., cough counts, durations) rather than open-ended signal exploration.

This limitation necessitates precise endpoint planning in trial design. Researchers must know in advance what features they will collect, how those features are derived, and what statistical assumptions underlie them. This makes edge-based approaches less suited for exploratory or hypothesis-generating studies, and more appropriate for trials with clearly defined mechanistic questions or validated digital biomarkers.

Balancing Privacy, Performance, and Complexity

Ultimately, the adoption of edge AI in clinical trials reflects a broader tension between privacy preservation and analytical richness. On-device models offer compelling benefits – reduced participant

burden, enhanced trust, and minimal data exposure – which come with constraints in flexibility, adaptability, and post hoc review⁴⁰.

Trial designers must weigh these tradeoffs in light of study goals, patient populations, and regulatory expectations. For studies where real-time monitoring, patient comfort, and ethical transparency are paramount – such as in chronic respiratory disease, sensitive therapeutic areas or any study that requires long term, continuous, longitudinal data – edge AI is the most viable and scalable solution, even if it requires a more structured and disciplined design approach.

References

1. A Turner, Richard Douglas, A Birring, Surinder S., Measuring cough: what really matters? *Journal of Thoracic Disease*, 2023, V 15, N 4, 2288–2299, 2077–6624, <https://jtd.amegroups.org/article/view/74803>
2. ACRO White Paper. Decentralizing Clinical Trials. A New Quality-by-Design, Risk-Based Framework (2020) – <https://www.acrohealth.org/resource/dctwhitepaper/>
3. Adiyaman A, Aksoy I, Deinum J, Staessen JA, Thien T. Influence of the hospital environment and presence of the physician on the white-coat effect. *J Hypertens*. 2015 Nov;33(11):2245–9. doi: 10.1097/HJH.0000000000000691. PMID: 26259118.
4. Agata Ferretti, Effy Vayena. In the shadow of privacy: Overlooked ethical concerns in COVID-19 digital epidemiology, *Epidemics*, Volume 41, 2022, 100652, ISSN 1755–4365, <https://doi.org/10.1016/j.epidem.2022.100652>.
5. Andreas Nautsch, Abelino Jiménez, Amos Treiber, Jascha Kolberg, Catherine Jasserand, Els Kindt, Héctor Delgado, Massimiliano Todisco, Mohamed Amine Hmani, Aymen Mtibaa, Mohammed Ahmed Abdelraheem, Alberto Abad, Francisco Teixeira, Driss Matrouf, Marta Gomez-Barrero, Dijana Petrovska-Delacrétaz, Gérard Chollet, Nicholas

⁴⁰ De Brouwer et al. 2021 outlines how edge computing and federated learning can modernize clinical trials. Proposes that processing data on participants' own devices (with algorithms running locally) allows continuous collection of digital biomarkers while enhancing privacy compared to centralized methods.

6. Anita Jumai Ehidiamen and Olajumoke Omolayo Oladapo. Enhancing ethical standards in clinical trials: A deep dive into regulatory compliance, informed consent, and participant rights protection frameworks, *World Journal of Biology Pharmacy and Health Sciences*, 2024, 20(01), 309–320. DOI: 10.30574/wjbphs.2024.20.1.0788
7. Barochiner J, Marín MJ, Janson JJ, Conti PR, Martínez R, Micali G, Conte IE, Plazzotta F. White Coat Uncontrolled Hypertension in Teleconsultation: A New and Frequent Entity. *High Blood Press Cardiovasc Prev*. 2022 Mar;29(2):155–161. doi: 10.1007/s40292-021-00498-y. Epub 2021 Dec 14. PMID: 34905157; PMCID: PMC8669402.
8. Berkhout C, Berbra O, Favre J, Collins C, Calafiore M, Peremans L, Van Royen P. Defining and evaluating the Hawthorne effect in primary care, a systematic review and meta-analysis. *Front Med (Lausanne)*. 2022 Nov 8;9:1033486. doi: 10.3389/fmed.2022.1033486. PMID: 36425097; PMCID: PMC9679018.
9. Chaccour, C., Sánchez-Olivieri, I., Siegel, S. et al. Validation and accuracy of the Hyfe cough monitoring system: a multicenter clinical study. *Sci Rep* 15, 880 (2025). <https://doi.org/10.1038/s41598-025-85341-3>
10. Coravos A, Goldsack JC, Karlin DR, Nebeker C, Perakslis E, Zimmerman N, Erb MK. Digital Medicine: A Primer on Measurement. *Digit Biomark*. 2019 May 9;3(2):31–71. doi: 10.1159/000500413. PMID: 32095767; PMCID: PMC7015383.
11. De Brouwer W, Patel CJ, Manrai AK, Rodriguez-Chavez IR, Shah NR. Empowering clinical research in a decentralized world. *NPJ Digit Med*. 2021 Jul 1;4(1):102. doi: 10.1038/s41746-021-00473-w. PMID: 34211085; PMCID: PMC8249659.
12. De Jong AJ, van Rijssel TI, Zuidgeest MGP, van Thiel GJMW, Askin S, Fons-Martínez J, De Smedt T, de Boer A, Santa-Ana-Tellez Y, Gardarsdottir H; Trials@Home Consortium. Opportunities and Challenges for Decentralized Clinical Trials: European Regulators' Perspective. *Clin Pharmacol Ther*. 2022 Aug;112(2):344–352. doi: 1
13. Fagherazzi G, Bensoussan Y. The Imperative of Voice Data Collection in Clinical Trials. *Digit Biomark*. 2024 Nov 13;8(1):207–209. doi: 10.1159/000541456. PMID: 39539994; PMCID: PMC11560146.
14. Fan Chung, Peter Small, Carlos Chaccour. Longitudinal Cough Frequency Monitoring in Subjects with Persistent Cough: Daily Variability and

- Predictability Over 30 Days, *European Respiratory Journal* 2024 64(suppl 68): PA787; DOI: <https://doi.org/10.1183/13993003.congress-2024.PA787>
15. Fiske, A., Henningsen, P., & Buyx, A. (2019). Your Robot Therapist Will See You Now: Ethical Implications of Embodied Artificial Intelligence in Psychiatry, Psychology, and Psychotherapy. *Journal of Medical Internet Research*, 21(5), e13216
 16. Galvosas M, Gabaldón-Figueira JC, Keen EM, et al. Performance evaluation of the smartphone-based AI cough monitoring app – Hyfe Cough Tracker against solicited respiratory sounds. *F1000Res*. 2023 Jun 9;11:730. doi: 10.12688/f1000research.122597.1. PMID: 39931660; PMCID: PMC11809693.
 17. Harrer, S. et al. Artificial intelligence for clinical trial design. *Trends in Pharmacological Sciences* (2019), 40(8), 577–591.
 18. Henry SG, Jerant A, Iosif AM, Feldman MD, Cipri C, Kravitz RL. Analysis of threats to research validity introduced by audio recording clinic visits: Selection bias, Hawthorne effect, both, or neither? *Patient Educ Couns*. 2015 Jul;98(7):849–56. doi: 10.1016/j.pec.2015.03.006. Epub 2015 Mar 17. PMID: 25837372; PMCID: PMC
 19. Isola S, Al Khalili Y. Protected Health Information. [Updated 2023 Jan 30]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan-. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK553131/>
 20. Kostick, K. M. et al. Behavioral health clinical trials and the ethics of digital phenotyping. *Journal of Medical Ethics* (2019), 45(12), 758–764.
 21. Kotevska, O. , Johnson, J. and Kusne, A. (2022), Analyzing Data Privacy for Edge Systems, *IEEE International Conference on Smart Computing*, Espoo, FI, [online], https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=934685 (Accessed April 22, 2025)
 22. M. Heidelberg, A. Kelman, J. Hopkins, M.E. Allen. The evolution of data ethics in clinical research and drug development. *Ethics, Medicine and Public Health*, Volume 14, 2020, 100517, ISSN 2352–5525, <https://doi.org/10.1016/j.jemep.2020.100517>.
 23. Macrae, Duncan. (2007). The Council for International Organizations and Medical Sciences (CIOMS) Guidelines on Ethics of Clinical Trials. *Proceedings of the American Thoracic Society*. 4. 176–8, discussion 178. 10.1513/pats.200701-011GC.

24. Narayanasetty, Swaroop & Jallu, Dr. (2021). A Review on Virtual Clinical Trials: The Future. *International Journal of Pharmaceutical Sciences Review and Research*. 68. 10.47583/ijpsrr.2021.v68i01.019.
25. Nass SJ, Levit LA, Gostin LO, editors. *Institute of Medicine (US) Committee on Health Research and the Privacy of Health Information: The HIPAA Privacy Rule; Beyond the HIPAA Privacy Rule: Enhancing Privacy, Improving Health Through Research*. Washington (DC): National Academies Press (US); 2009. 2, The Value and Import
26. Nebeker C, Harlow J, Espinoza Giacinto R, Orozco-Linares R, Bloss CS, Weibel N. Ethical and regulatory challenges of research using pervasive sensing and other emerging technologies: IRB perspectives. *AJOB Empir Bioeth*. 2017 Oct-Dec;8(4):266–276. doi: 10.1080/23294515.2017.1403980. PMID: 29125425.
27. Nebeker C, Torous J, Bartlett Ellis RJ. Building the case for actionable ethics in digital health research supported by artificial intelligence. *BMC Med*. 2019 Jul 17;17(1):137. doi: 10.1186/s12916-019-1377-7. PMID: 31311535; PMCID: PMC6636063.
28. Nebeker, C. et al. Ethical and regulatory challenges of research using pervasive sensing and other emerging technologies: IRB perspectives. *AJOB Empirical Bioethics*, (2017) 8(4), 266–276.
29. Nuredini G, Saunders A, Rajkumar C, Okorie M. Current status of white coat hypertension: where are we? *Ther Adv Cardiovasc Dis*. 2020 Jan-Dec;14:1753944720931637. doi: 10.1177/1753944720931637. PMID: 32580646; PMCID: PMC7318827.
30. Office for Human Research Protections (OHRP) – *Guidance on Research Involving Coded Private Information or Biological Specimens*. (2018)
31. Ostherr, Kirsten & Borodina, Svetlana & Bracken, Rachel & Lotterman, Charles & Storer, Eliot. (2017). Trust and privacy in the context of user-generated health data. *Big Data & Society*. 4. 10.1177/2053951717704673.
32. Reddy, K.D.; Chawla, S. Wearable Technology in Gastroenterology: Current Applications and Future Directions. *J. Clin. Med*. 2025, 14, 2403. <https://doi.org/10.3390/jcm14072403>
33. Tackney MS, Steele A, Newman J, Fritzsche MC, Lucivero F, Khadjesari Z, Lynch J, Abbott RA, Barber VS, Carpenter JR, Copsey B, Davies EH, Dixon WG, Fox L, González J, Griffiths J, Hinchliffe CHL, Kolanko MA, McGagh D,

- Rodriguez A, Roussos G, So KBE, Stanton L, Toshner M, Varian F, Williamson PR, Yimer BB, Villar SS. D
34. Tanner RM, Shimbo D, Seals SR, Reynolds K, Bowling CB, Ogedegbe G, Muntner P. White-Coat Effect Among Older Adults: Data From the Jackson Heart Study. *J Clin Hypertens (Greenwich)*. 2016 Feb;18(2):139-45. doi: 10.1111/jch.12644. Epub 2015 Aug 17. PMID: 26279070; PMCID: PMC4742426.
35. Topol, E. J. High-performance medicine: the convergence of human and artificial intelligence (2019). *Nat Med*, 25, 44-56.
36. van Rijssel, T.I., van Thiel, G.J.M.W., van Delden, J.J.M. et al. The Ethics of Decentralized Clinical Trials and Informed Consent: Taking Technologies' Soft Impacts into Account. *Health Care Anal* (2024). <https://doi.org/10.1007/s10728-024-00483-1>
37. Yusuf ZK, Dixon WG, Sharp C, Cook L, Holm S, Sanders C. Building and Sustaining Public Trust in Health Data Sharing for Musculoskeletal Research: Semistructured Interview and Focus Group Study. *J Med Internet Res* 2024; 26:e53024. doi: 10.2196/53024