AMBIENT DIGITAL SCRIBE: CONSIDERATIONS FOR USE IN CLINICAL TRIALS

An industry analysis of ambient digital scribe (ADS) technology use in clinical research and in the development of new pharmaceutical and medical device products.

Elise Felicione, Jeff Lee, Jonathan Helfgott, Craig Serra, Craig Lipset, Raj Ratwani, Petros Okubagzi, Bert Hartog

August, 2025

Prelude

New technologies enable audio recording of interactions between clinicians and patients, followed by the conversion of speech to text and the analysis, structuring, and translation of the transcribed information by medical conversation-trained generative artificial intelligence (Gen-AI) large language models (LLMs).

The outputs of this process include items typically created manually by clinicians, such as Subjective, Objective, Assessment, and Plan (SOAP) notes, a patient visit summary written at an accessible reading level in the patient's native language, as well as data organized for integration with electronic health record systems to update a patient's medical record.

When data aggregation and synthesis models are added to the mix, these systems function as "clinician co-pilots," drafting a wider array of clinical documents, including pre-visit preparation and integrated clinical summaries for patient referral or hospital shift changes.

These technologies, known as **Ambient Digital Scribes (ADS)**, are being rapidly adopted by healthcare providers, clinics, and health systems in the United States and worldwide. They have been shown to reduce clinician time spent on computers, alleviate burnout and task load associated with clinical documentation, automate medical coding and billing, and improve the patient and clinician experiences by enabling clinicians to focus more on the conversation and less on typing during visits.

ADSs are creating new norms for clinician documentation and medical record generation. Keyboards in exam rooms may become a relic of the past.

Given the potential for ADSs to enhance the overall conduct of research, it will be crucial for our industry to establish pathways for the responsible use of these technologies in research.

Executive Summary

This paper explores the ways ADS can impact, or already may be impacting, clinical research and trials, addressing the topic from several angles:

1. Use of these technologies, designed for care delivery, out-of-the-box to record and document clinical trial encounters, along with the associated benefits and risks of this practice.

2. Ways ADS creates unprecedented pathways for trial data generation and how they can be leveraged to acquire data, detect safety signals, enhance quality, and identify eligible patients for clinical trials.

3. Ways these technologies may need to be adapted to contemplate clinical research's unique needs and requirements, including operational and implementation considerations.

4. Metrics for performance and impact of ADS in the context of research, and the outcomes that matter most.

5. Clinical trial regulations and guidelines, AI in drug development guidelines, standards for the responsible use of AI, and ICH GCP to guide the industry's responsible and compliant use of ADSs in clinical trials. It includes a proposed GCP inspection readiness checklist that can be used by sponsors, contract research organizations (CROs), and sites to ensure that ADSs are fit for trial data generation in the context of an FDA- or other health authority-regulated clinical trial.

- **Intended Audience:** professionals and leaders in industries spanning clinical research, healthcare, clinical trial software development, and ADS solution developers. This content will likely be new territory for readers in at least one aspect, given that the ADS and clinical trial fields have not yet intersected.

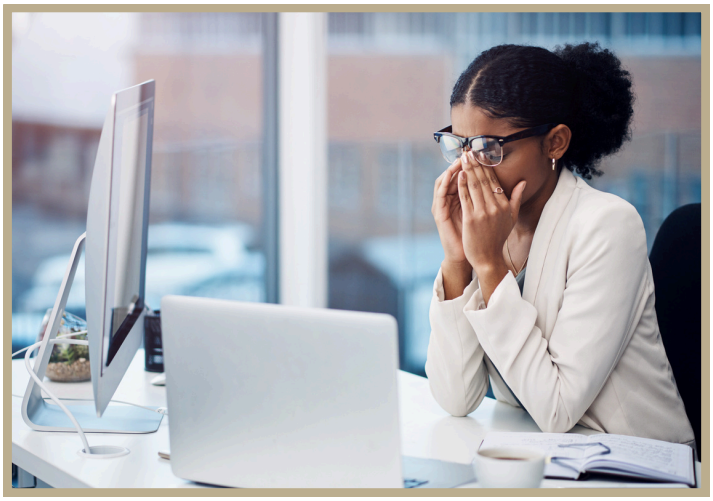
- **Authors' Aim**: to disseminate the prompt for thought on this topic, invite public comments, additions, and opinions, and inspire actions to enable ADS use in clinical research and trials.

Table Of Contents

Background

The Epidemic of Clinician Burnout

Professional burnout marks the beginning of the ADS story. Burnout is a psychosomatic stress response characterized by emotional exhaustion, depersonalization, and a diminished sense of accomplishment (1,2). It is especially prevalent among clinicians and has reached crisis levels globally, with 46% of clinicians reporting feeling symptoms of burnout and 44% intending to look for a new job (3). In 2022, 71,000 physicians quit the profession in the United States, resulting in $4.6 billion in attributed turnover costs. A shortage of 187,130 physicians, or 16% of the workforce, is projected for 2037 in the United States (4). A recent study found that primary care physicians require an average of an impossible 26.7 hours daily to keep up with their workload (5).



Broad calls to action have been made to stakeholders across the healthcare ecosystem to address provider burnout. Research into causes and opportunities for the prevention and treatment of burnout reveals the burden of clinical documentation, particularly the use of the electronic medical record (EMR) system, as a significant source of stress (7). For every hour of patient care delivered, the clinician spends nearly 2 hours in the EMR system (8,9).

And not all of this work can be accomplished during work hours at the clinic: one in five clinicians reports spending 8 hours per week catching up on medical record maintenance during non-work hours, known casually as “pajama charting” (10). Recommended mitigations to address this stress include having a medical assistant take notes during visits.

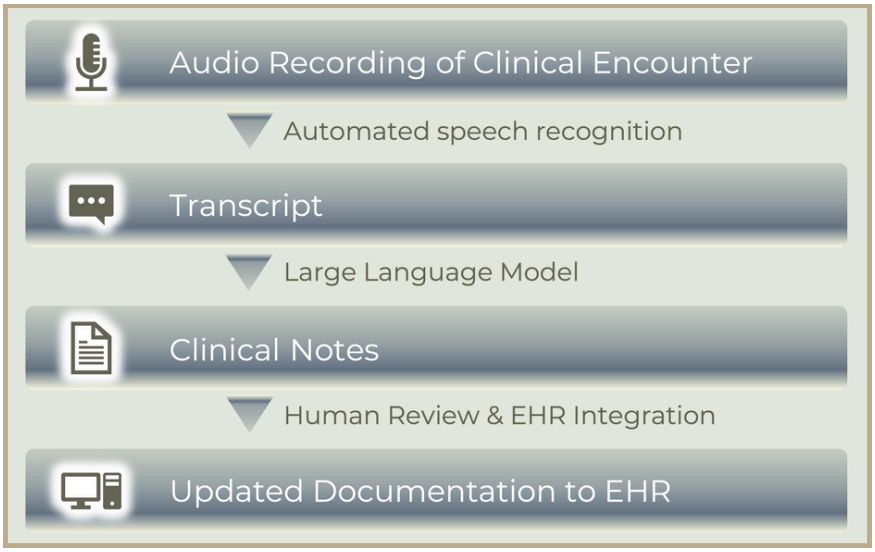
The Rise of The Ambient Digital Scribe

Just as clinician burnout reached crisis levels in 2022, another force in healthcare was on the rise: Gen-AI and machine learning. This created a natural alignment of problem and solution: a machine to draft notes and documentation for clinicians.

Entrepreneurs and tech companies, both large and small, have stepped up to create what today accounts for at least 88 solutions (11). Figure 1 provides a visual representation of how these tools operate.

After the patient consents to recording the visit and agrees to the terms of use, the clinician turns on the internal microphone on the exam room computer or a mobile device. An audio recording of the entire clinician-patient exchange is created. A transcript is created using phonetic analysis, language modeling, and medical lexicon integration. Next, large language models (LLMs) interpret the transcript. The raw transcript is parsed into individual components such as symptoms, diagnoses, and medications.

Figure 1.



Adapted from Elion AI Clinician Copilots: Technology, Market Trends, and Key Differentiators (Jan 23, 2025).

The LLM then inserts these components into pre-defined, audience-specific templates, creating documents typically produced by clinicians, such as SOAP notes and plain language after-visit patient summaries, as well as coding of procedures and diagnoses for billing. The clinician retains full editorial rights and can add, delete, or correct any AI-generated content. Eventually, the clinician approves the content and uploads it into the EMR system through integrations or manual cut-and-paste. (Newer, deeply integrated systems enable the performance of these tasks directly in the EMR.)

Health systems have introduced these tools into their ecosystems, often through progressive quality improvement research programs and pilots (12) that sequentially derisk the use of ADS technologies: first, in mock environments (13), then parallel tool use and manual note taking, then single clinic implementation, broader clinic and specialty implementations, and finally, full-scale enterprise adoption.

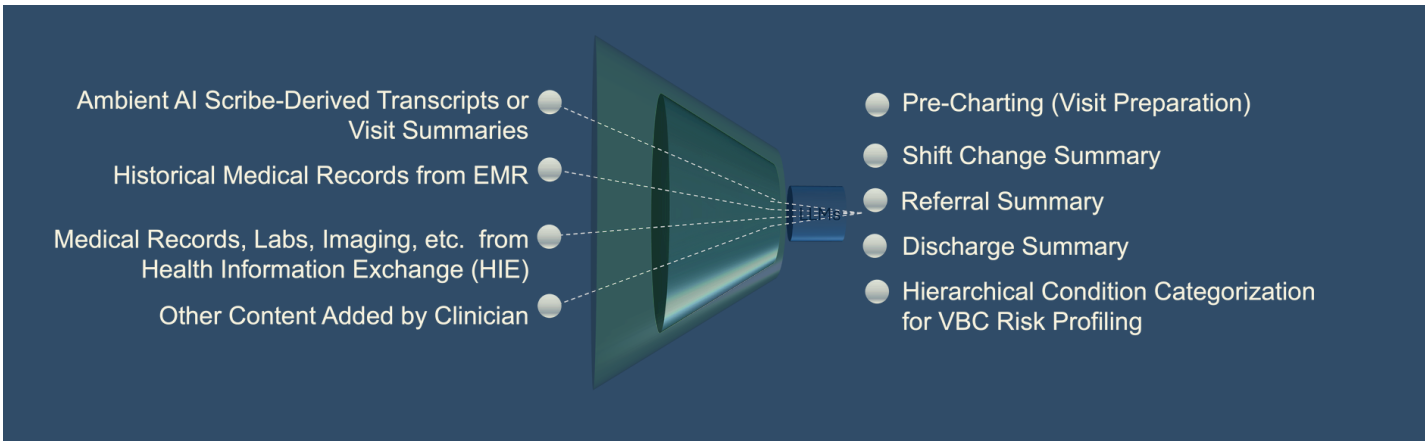
The Rapid Adoption of Ambient Digital Scribes in Healthcare

Early adopters have published pilot results in peer-reviewed journals, showing positive performance indicators, such as reducing time at the computer by one hour per day; improvement of burnout, dis-engagement, and task load measured with validated psychometric instruments; favorable survey results about ease of use, perceived efficiency, and documentation quality, and that it improves the patient-physician interaction. “It makes the visit so much more enjoyable because now you can talk more with the patient and concentrate on their concerns,” said one participating physician (14, 15, 16, 17, 18). It should be noted, however, that patient experience evidence is lacking (19).

Fierce Healthcare interviewed analysts who predict that by the end of 2025, 30% of the healthcare market will utilize AI scribes, and 40-50% of health systems will implement some form of AI scribing technology (20). The Wall Street Journal reports that these technologies are moving into inpatient and emergency settings as well (21). Analysts anticipate that the field could narrow significantly, predicting that the larger health systems market will consolidate around solutions that integrate with their EMRs. In August 2025, Epic announced plans to release its own Microsoft-powered scribe with embedded features in its EMR systems (22).

Pressure is mounting to deliver accuracy and personalization for individuals or specialties. Elion predicts a broadening of the functionality of ADSs into so-called “AI Clinician Co-Pilots” (23), which will, along with documenting single encounters, create integrated clinical summaries by aggregating medical records and histories from multiple sources, as depicted in Figure 2. Clinical decision support features are also in development, with the understanding that they do not override the clinician. “This is a co-pilot, not a pilot or an autopilot, and the physician is always in command of the situation,” said Joe Petro, Vice President of Health and Life Sciences for Microsoft.

Figure 2. AI Clinician Copilots integrate health information to produce customized clinical summaries.



Benefits, Risks, Gaps: ADS in Clinical Research

As ADSs become more common in healthcare and considering implementation strategies that encourage consistent workflows and routines, it is reasonable to assume that investigators who use scribes as part of care delivery may automatically turn on the recording device and conduct a trial visit using these technologies.

Benefits of using healthcare ADS technologies in clinical trials

ADS can offer benefits to research investigators similar to those cited by clinicians. Although investigator burnout is discussed less frequently than physician burnout, it is common knowledge that one in every two investigators leaves research after conducting just one trial, due to workload, competing priorities, staff time demands, data and safety reporting requirements, and a lack of financial support and recognition (24). The burden of clinical research is further intensified by increasing protocol complexity and a 300% rise in data points collected during pivotal trials from 2010 to 2020 (25).

In clinical trials, it is often the clinical research coordinators who execute tasks related to documentation, and much more is published about workload and attrition. There is a critical shortage of clinical research coordinators and nurses, with a 7-10x ratio of vacancies to job seekers, and research staff turnover in site networks is estimated to be 35-61% (26). Study coordinators report feeling stress due to sponsor-imposed timelines, protocol amendments, long work hours, budget constraints, data entry requirements, and document maintenance and submission to sponsors (27). One in five coordinators surveyed in an Italian study reported having significant workload-related stress (28).

The burden associated with clinical trial documentation differs from that of clinical care. On the one hand, data origination is more straightforward: the site conducts the assessments precisely as outlined in the protocol; data are inherently objective, and in cases of subjectivity, the protocol provides extensive guidance to ensure consistency across multiple sites. On the other hand, data must be aggregated from numerous sources, such as central labs or imaging, and external medical records, including lengthy or indefinite medical histories in some cases. Reporting is scrutinized by monitors and subject to audit and inspection, requiring extra attention to be paid to accurate documentation compared to clinical care.

ADS can enhance the participant experience in clinical trials by allowing the researcher to step away from the chart and concentrate on the conversation. This can reduce the “guinea pig” stigma sometimes associated with research participation.

ADS can enhance the investigator—participant experience in clinical trials in similar fashion as it does in care delivery by allowing the clinician researcher to step away from the research chart and concentrate on the conversation. This can reduce the "guinea pig" stigma that can sometimes be linked to clinical trials and be a deterrent to enrollment and retention.

Risks of using healthcare ADS technologies in clinical trials

A common complaint among ADS technology users is **errors** (29), which fall into four categories:

- **Hallucinations:** AI records a clinically relevant finding that did not occur during the encounter; "false positive"
- **Omissions:** AI fails to note a clinically relevant finding that did occur; "false negative"
- **Misunderstandings:** AI inaccurately describes or characterizes a finding in a clinically meaningful way
- **Inclusion of irrelevant content**.

A study by Hose, *et. al*. found an average of 6 errors per visit where ADS was used, 45% of which were of moderate clinical significance, and 83% were errors of omission (30). Biro *et al*. found that when errors appeared in AI-generated draft responses to patient portal messages, approximately 75% of physicians failed to identify the error introduced by the generative AI (31). Errors of omission are particularly problematic as it's harder for a human to notice something that is not present than something that is erroneously present. Additionally, misunderstandings can introduce risks in that they are more easily overlooked.

The inclusion of irrelevant or repetitive content can lead to longer, less succinct notes that dilute meaningful and actionable items, potentially leading to their oversight.

This risk is mitigated in practice, as the investigator can review AI outputs and correct any errors before approving the data. Most systems also allow clinicians to access the original transcript to verify the verbatim conversation. Nevertheless, there remains a risk that something may be overlooked during the human review, potentially allowing mistakes to be carried forward into source documents, the Case Report Form (CRF), and the clinical trial database.

It should be noted, however, that the true rate of error with manual, human clinical data recording in the source is unknown and unverifiable, and cannot be presumed to be zero.

Additional Risks

Bias may be introduced when the model is trained and fine-tuned on data that does not represent the context of use. There may be terminology specific to certain specialties, languages, and figures of speech that could impact the model's accuracy and

detection abilities. A 2023 study demonstrated that the spoken phrases “um-hmm” and “uh-huh"—known as non-lexical conversational sounds—exhibited an astonishing 41-57% error rate in non-clinical conversations and a 95-99% error rate in clinical conversations (32).

Privacy and security are also risks. ADS systems are designed for HIPAA compliance and the equivalent privacy requirements in markets outside the United States. For Clinical Trials, there is the added requirement for research participant identity protection, and no personal identifiers may be transmitted outside of the site.

ADS technologies deployed in practice settings were designed for healthcare delivery and not for the specific key needs in research. For example, their language models may not be tuned for emerging research therapy areas such as rare diseases or cell- and gene-therapies.

Research may introduce new and different assessments that are not performed in clinical care and thus have not been used in training or tuning the LLMs. Furthermore, clinicians report that existing clinical scribe tools are less helpful for objective, or “checklist-driven,” visits (33), a characteristic of clinical trial visits.

ADS technologies deployed in practice settings were designed for healthcare delivery and not for the specific needs of research.

Finally, representativeness and bias is a risk. Clinical trial participants can differ from the general population from whom LLM training data was derived. In research, younger, health-literate and urban-dwelling individuals are often over-represented; older adults, women, mobility-challenged, and ethnic minorities are often underrepresented; trials include a narrowed subset of individuals that may have fewer comorbidities or are taking fewer or different medications. (34, 35)

Gaps and Unmet Needs

As noted, the LLM approaches will need further refinement to reduce error rates and meet the standards required for clinical trials. Patient-facing outputs, beyond simple clinical notes, will require IRB review and approval of their templates before they can be used.

Finally, the ultimate destination for content in a clinical trial often differs from that in healthcare. In healthcare, the destination is typically a patient’s chart or the Electronic Medical Record (EMR) system. In contrast, it is the source document for a clinical trial (which can be separate from the EMR), and ultimately the Case Report Form.

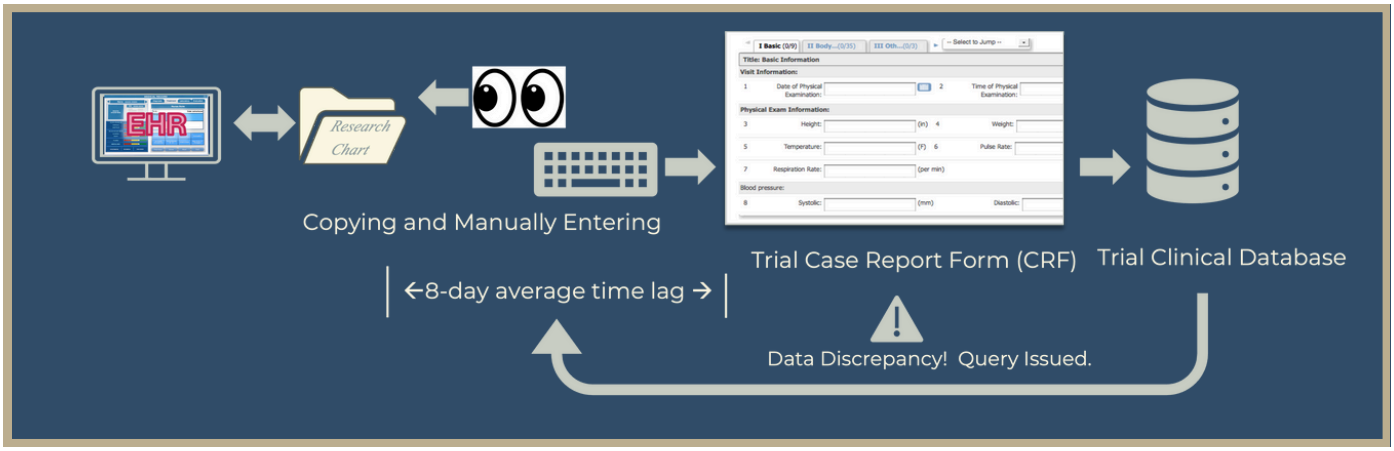
New Opportunities Offered by ADS

ADS Can Enable Direct Data Capture, Enhance Quality, Trial Recruitment and Retention, and Support Pharmacovigilance.

Direct Data Capture

Today, data acquisition from patient encounters often relies on human-written source data recording, followed by manual transcription into case report forms (CRFs). This sequential and manual process can result in data reporting lag time, transcription errors, and incomplete data (36), as described in Figure 3.

Figure 3.



After acquiring the data, the next step is quality control. The first line of quality control is performed by monitors employed by sponsors or contract research organizations (CROs), who typically visit sites every 4 to 8 weeks to conduct source document reviews (SDR) and source data verification (SDV). It is estimated that 25-30% of a trial's total budget is allocated for monitoring. Monitors report spending 46% of their monitoring time on SDV (37), resulting in direct costs of $5-10M in a late-development clinical trial, alongside additional direct and indirect costs, including site fees for monitoring visits that range from hundreds to thousands of dollars per monitor-day on-site. Next comes auditing, and finally, patient and investigator site records from the trial are subject to inspections by health authorities.

Direct Data Capture, cont.

Over the past thirty years, significant innovations have occurred to enhance and automate data acquisition and quality control. Analytical, risk-based monitoring (ARBM), adaptive, and remote monitoring methods enable the targeted utilization of monitoring resources on critical data points in real-time, based on various risk indicators throughout the study (38). Technology has also transformed data management. Electronic Case Report Forms (eCRFs) in electronic data capture systems (eDC) are essential in clinical trials today. Additionally, electronic Clinical Outcomes Assessment (eCOA), electronic source documentation (eSource), and electronic consent form systems (eCRF) have become prevalent in trials and have integrated into multi-solution platforms.

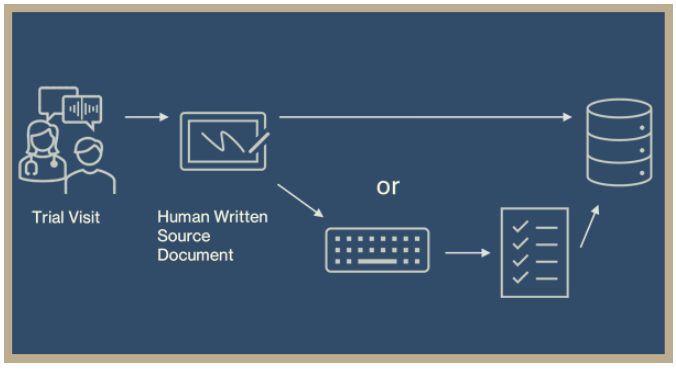
The technological transformation enables the direct acquisition of trial data from electronic systems without the need for human transcription, commonly referred to as eSource to eDC. However, despite extensive efforts and investment, transcription-free data transfer from eSource to eDC has struggled to achieve scale. EMRs, which many health systems prefer as their source data system, require custom configuration to map data, even with advances in standardized HL7 data structure and new FHIR/VULCAN APIs. Many EMRs globally do not meet system qualifications and requirements outlined by health authorities and thus cannot be used as eSource.

By-design eSource software systems integrate data interoperability and regulation compliance into their core design, but often only integrate with EDC systems on the same platform. Sites utilizing EMRs consider separate eSource systems redundant or too expensive, or researchers prefer to create paper source documents ("research charts") that can serve as a basis for manual transcription into EMR and EDC.

How ADS Could Break the Barrier to eSource At Scale

Today, when the investigator hand-writes or types a data point from a clinical encounter in a source document, this is "time zero" in its lifecycle. It requires manual entry into the source document (either paper or electronic). From there, the data reaches the trial database in one of two ways: (1) it is retrieved digitally by leveraging interoperability protocols or (2) it is manually copied or transcribed into a Case Report Form to transfer the data into the trial database, as shown in Figure 4.

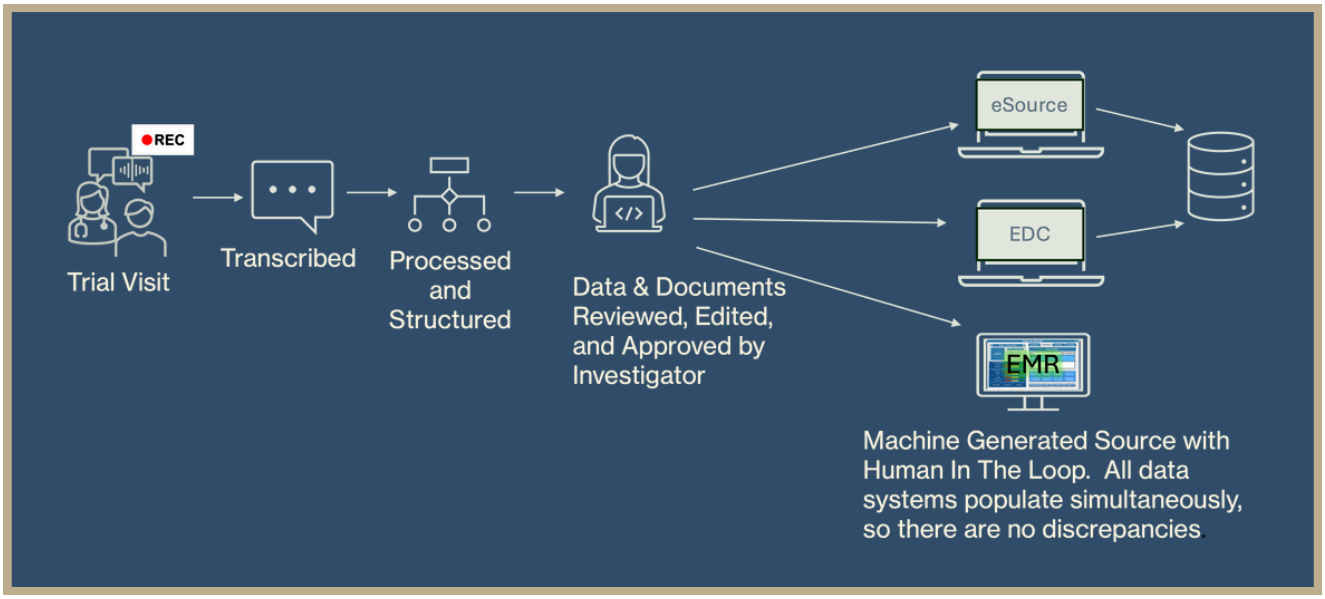
Figure 4.



How ADS Could Break the Barrier to eSource At Scale

With ADS, however, “time zero” for a data point occurs upstream of the source document. One could even think of it as the moment the sound wave flows through the exam room and is picked up by an audio recording device. Or the point at which that audio recording is converted into a transcript. Either way, the origin of data for live encounters has an unprecedented first step. With it comes an unprecedented opportunity to capture the data earlier and transfer it to the case report form or the trial database. Figure 5 describes how trial encounters could be documented using ADS.

Figure 5.



The trial visit is recorded and transcribed, with the data processed and structured. The investigator or delegate will then review the data for accuracy and completeness, referring to the transcript as needed during this review. Once the investigator is satisfied with the content, they will approve and publish it. Publishing means the data are simultaneously pushed to the clinical trial systems and their EMR system. After the data are published and used to populate systems, the system designated by the investigator as the source will be considered the "Source" and will be subject to handling according to regulations and guidelines for source data. The virtue of this schema are that it circumvents barriers to passively obtaining data and avoids manual transcription.

ADS For Quality

ADS could track whether the protocol-required procedures are completed and alert the researcher if any are missed. It could detect needs for investigator and staff retraining on the protocol, potential protocol deviations or discrepancies, or other quality risk areas, enabling corrections to avoid errors. Additionally, it can be used for training in bedside manner or conducting the informed consent process.

ADS For Pharmacovigilance

ADS offers passive pathways to detect adverse events or safety signals. It could alert the investigator and staff to collect follow up information, flag an event as time sensitive for reporting if potentially a serious event and/or unexpected event, and generate safety reports in structure or content to be relayed to IRBs/IECs and the sponsor's pharmacovigilance teams, as required by the protocol.

ADS For Recruitment and Retention

ADS could also provide a new method for identifying patients who qualify for clinical trials during healthcare encounters. The audio transcript could be compared with a trial's inclusion and exclusion criteria to determine eligibility. The provider could be alerted to the trial for which the patient may qualify. The transcript provides unprecedented insights for patient retention and engagement as well. For instance, if a patient mentions difficulties with traffic or finding childcare, the site can be alerted to opportunities to assist with transportation or to offer home or community health services related to the trial.

ADS For Clinician Reported Outcome Assessment

Clinician Reported Outcome Assessments could be completed in the background during an encounter, subject to clinician review and verification of the information.

Key Adaptations to ADS Technologies Required for Use in Clinical Trials

The earlier section notes risks and gaps with ADS technologies used to record clinical research encounters. To make this technology suitable for clinical trials, adaptations may be necessary in six key areas.

1. **Regulatory Review and Clearance.** Existing ADS technologies used in clinical care have not undergone the same level of regulatory review that may be necessary for AI systems employed in drug development. Insights into regulators' expectations are beginning to emerge, suggesting that these systems may need validation and clearance from regulators before clinical trial data supporting market authorization applications can be generated using ADSs. (More details on regulatory considerations will follow.)

2. **Accuracy & Completeness Rate.** The system's accuracy rate may require improvement. It must be precise for critical data, such as endpoints, medications, and adverse events. The exact quality limits for clinical trials should be determined through discussions with quality experts and regulatory officials. However, a general benchmark should be that the combined output of the AI and human reviewers is judged as at least non-inferior to what would be expected if the data were originated entirely by humans, as it is currently done. This could be tested before deployment through dual-documentation and independent comparison for accuracy and completeness.

3. **New Data Standards.** The ADS technology must be capable of extracting data from a transcript and structuring it to either auto-populate a trial case report form or conform to clinical trial data standards, such as Clinical Data Interchange Standards Consortium (CDISC). The large language model must be trained to accurately map the transcript's information to the appropriate trial data elements and render them according to foundational and data exchange conventions for research.

4. **New Integrations.** Integrations should be built to enable data ingestion by the major EDC and eSource systems. This is particularly important as eSource systems are adopted by many research-exclusive sites and site networks in similar fashion as health systems have adopted EMR systems. EMR integrations are a key factor in health system ADS adoption, and sites using eSource systems will likely expect system integrations. Integrations of technology and workflows should be explored in partnership with these companies.

Key Adaptations to ADS Technologies Required for Use in Clinical Trials

The earlier section notes risks and gaps with ADS technologies used to record clinical research encounters. To make this technology suitable for clinical trials, adaptations may be necessary in six key areas.

5 **General Research and Protocol-Specific Configuration.** The visit transcript should be cross-referenced with the clinical trial protocol to ensure the system understands which assessments will be performed and how they should be conducted. In healthcare, the ADS system is aware of basic details about the type of visit and uses background contextual information as input to assist in interpreting the visit transcript. The same will need to apply to use in clinical trials.

6 **Research Privacy and Security.** The system may require adaptations to comply with local privacy laws, data security regulations, and research regulations, which will be discussed in greater detail later.

Operational and Implementation Considerations

The impact of ADS used for clinical research—encompassing people, processes, and technology—must be well understood by adopting institutions. Barriers and facilitators will need to be carefully assessed, and stakeholders will need to participate in assessing fit, planning and executing implementation, and measuring performance. This is best explored through pilot tests designed to learn how to implement ADS effectively in research. An implementation team should be established, and appropriate resources should be provided to guide the implementation and measure performance metrics along the way.

Health systems that have implemented ADS for care delivery are unlikely to adopt a new system for research; instead, they will expect to use their existing system without disrupting established data flows to the EMR. Research-exclusive institutions will view ADS as a new entrant in their ecosystems and will be concerned about the burden and risks associated with its adoption. Adoption will be encouraged if ADS functionalities are integrated as features in existing systems that sites already use or are mandated to use by sponsors, such as EDC or eSource systems. Finally, use of ADS in a research study will be optional: countries may not allow it; sites are free to choose to use it or not; and patients may decline to give consent. Therefore, fully manual documentation and data acquisition processes must coexist with processes including ADS.

Measuring Impact and Outcomes

Measuring Outcomes in Healthcare Practice

Ambient Digital Scribes have been considered one of the most quickly adopted technologies in the history of health tech, bypassing the more customary approach of in depth strategic planning, rigorous testing, evidence generation of system performance, and measurement of outcomes, such as employee retention, financial metrics, patient retention, and quality of medical records and care delivery.

In September 2024, the Peterson Health Technology Institute (PHTI) convened a diverse panel of 63 stakeholders representing health system purchasers and users, ADS technology innovators, investors, and other experts, who performed a comprehensive assessment of real-world use of ADSs, how adoption decisions are being made, and developed a standard assessment framework to guide future decisions on adoption and sustainment of ADSs in practice (39). The report found significant heterogeneity in measurement approaches used across health systems and gaps in longitudinal evidence due to the novelty of these technologies. They developed a consolidated metrics framework, as shown in Figure 6.

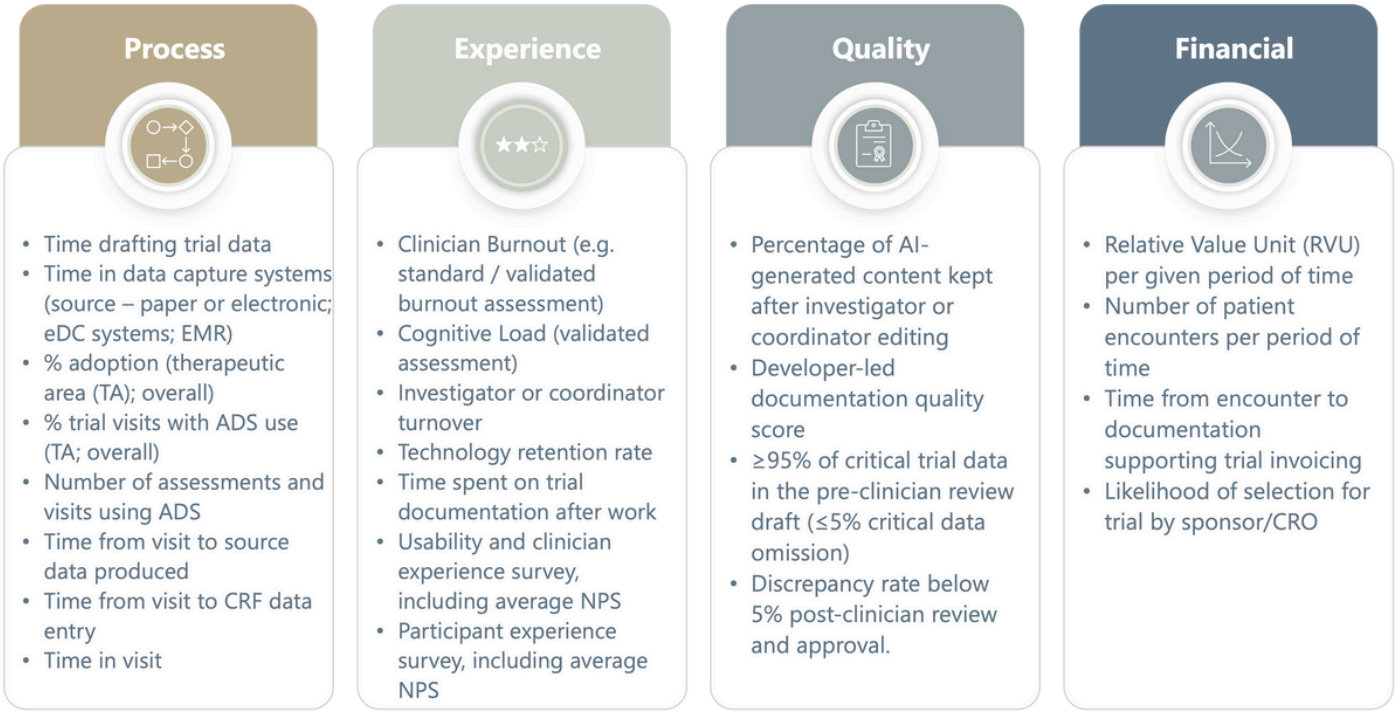
Figure 6. PHTI framework and recommended metrics and measurement approaches for evaluating ADS technologies in healthcare

Measuring Outcomes in Clinical Research

For clinical research, the healthcare framework can serve as a guide to think about what metrics and outcomes are meaningful, while being mindful that the aim of clinical research differs from that of healthcare. Clinical research aims to test a hypothesis; clinical trials aim to to generate evidence of an intervention's efficacy, safety, tolerability, usability, and/or real-world effectiveness to inform regulators, providers, patients, and payers about the trade-off of benefits vs. risks, and to support their decision as to whether or not to authorize marketing, prescribe, use, or cover, respectively. This said, the research process will not happen unless research is doable and carries sufficient relative-priority for all stakeholders – sponsors, sites, investigators, and patients - to other activities that compete with time and resources.

Keeping in mind the aim of clinical research – to generate evidence for decision support – as well as the need for participation from required stakeholders, which depends on feasibility and relative priority, we propose an adapted measurement framework specific to clinical research, as described in Figure 7.

Figure 7. Adaptation of the PHTI framework and recommended metrics in the context of clinical trials



Regulatory and Ethics Considerations

ADS is a new data generation modality that is not currently contemplated in regulations or guidelines. This section analyzes ADS's use in investigational product research. It proposes interpretations that should be further discussed, validated, and refined as necessary through peer review and direct engagement with regulators, policymakers, industry organizations, communities of practice, and other stakeholders. We present these interpretations solely to initiate a conversation with stakeholders.

Source Data Requirements and Principles

As discussed earlier, ADSs create a preliminary electronic record of information, subject to the clinician's review, editing, and approval. This said, many questions arise: would regulators view the transcript or machine-generated content before investigator review as "source," and, as a result, would these outputs be subject to regulations governing source and electronic source, including maintaining an audit trail and subjection to monitoring, audit, and inspection (40, 41, 42)? After all, it technically represents the first written record of the encounter, even if not written by a human.

Regulations require that source documents be "ALCOA+C" which stands for attributable, legible, contemporaneous, original, accurate, and complete.

The latter two criteria, accuracy and completeness, require the human in the loop to verify; therefore, documents that the investigator has not signed off on would likely not be suitable as the source.

Healthcare institutions, faced with similar questions about what constitutes a legal record, are encouraged to establish written policies in advance of record creation to clarify that transcripts and draft renditions of documentation are considered drafts and are not considered a legal record until the clinician has signed off on them (43). Such institutional policies could be broadened to cover clinical trial records as well.

Traditionally, clinical protocols require that information undergo investigator review and approval before it is considered "source." It also depends on the specifics of how ADS systems operate. For instance, if the information in the ADS is "locked" after being approved and transferred to a source data system, or a paper Source Document is published, the expectation is that further review or edits will occur in that external Source Document or eSource system, and not in the viewer/editor that comes with the ADS system.

One could view it as shorthand on a notepad that an assistant takes while the provider conducts the visit: the shorthand note will then be used by the investigator or delegate who actually conducted the assessment to create original notes.

Alternatively, it could be viewed as analogous to a draft document, right before the author designates it as “final” and assigns a version date. Earlier drafts of these documents, which are not shared beyond the authors, are not customarily considered official.

The same principle applies when ADS-derived visit data are used to identify eligible trial participants. This information, captured and in human-readable drafts, would be considered a draft and require subsequent cognitive processing and review from a qualified investigator and staff, resulting in an entry in a trial participant screening log or a similar document.

eSource System Regulations and Guidance for Industry

TransCelerate, having interpreted and distilled global regulations and guidance documents, categorizes eSource into four groups, noting that this may evolve with technological advancements (44): (1) Electronic Health Records; (2) Devices and Apps; (3) Non-CRF data that flows directly to the clinical database (e.g., central laboratory data); (4) Direct data capture (the direct entry of data by site staff into a mobile application or EDC system).

Naturally, the question arises: into which category would ADS-generated data fall? Is the AI functioning more like a blood-pressure cuff or an app? Or is it merely an assistant to expedite number 4 above?

We postulate that it is the latter: the AI assists the investigator by generating a draft. As discussed previously, the ADS functions similarly to the human assistant that helps the investigator or coordinator take notes during an encounter. The investigator, not the assistant, would be considered the originator of the data.

Implications of ADS In the Context Of ICH GCP Principles

To frame the consideration of ethical and compliance principles, we examine ADS alongside the principles of the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use, specifically Good Clinical Practices (ICH GCP), the guidelines upon which regional and country-specific laws and regulations governing human participants research are formed, in Table 1.

Table 1: ICH GCP Principles and Considerations for ADS Use

Principles of ICH GCP (No. and Principal) (45)	Considerations for ADS Use in Trials
1. Abide by the Declaration of Helsinki Principles (46)	**Respect for Persons:** individuals should be free to participate in the trial if they don't want their visit to be recorded. ADS should be used first and foremost to enhance their experience. **Beneficence and Non-Maleficence**: ADS should not cause incremental harm or risk to participants. **Justice and the protection of vulnerable individuals**: ADS should not disproportionately benefit or burden vulnerable individuals, and should consider the specific needs of these populations. For example, include child assent for minors; not result in the exclusion of hearing or speech-impaired individuals from the trial just for the sake of ease of visit recording. Also, use of ADS may not be appropriate for some vulnerable populations or for visits when sensitive topics will be discussed (e.g. substance abuse, domestic violence, or other unlawful activities).
2. Informed Consent	Participants must be able to give and withdraw their consent for ADS use without impact on their ability to join or continue in the clinical trial. Informed consent should be explicitly obtained related to the use of ADS, just as it is required in healthcare. It should be clear and concise, covering topics such as length of time that the recording will be stored, who can access the recording, how the investigator will review it, and whether any of the recording, transcript, physician edits, or final notes will be used to train the AI. It should cover risks (e.g. data loss, re-identification, interception, or error), be objective about benefits, and be written at appropriate reading level and in the participant's native language. How its use could impact healthcare coverage must be covered.

Table 1: ICH GCP Principles and Considerations for ADS Use (Cont.)

Principles of ICH GCP (Number and Principal)	Considerations for ADS Use in Trials
3. IRB / IEC Review	The use of ADS should be explicitly disclosed to a qualified institutional review board or independent ethics committee as required by law in the geographic area where it will be used. This should include how or when it will be used. The informed consent or any other participant-facing materials that explain the use of ADS in the trial should be reviewed and approved by the IRB/IEC, with the format, content, and layout with which it will be presented to the participant.
4. Scientific Soundness	Use of ADS in clinical trials should in no way alter the scientific validity of the research protocol.
5. Qualified Individuals	Practitioners involved in trial visits should be restricted to the investigator and qualified staff, who have documented evidence of required licensure, and who have been trained on the trial protocol and any IRB-approved amendments. ADSs should accurately attribute persons involved in a recorded encounter.
6. Quality By Design	Quality controls and assurance should be implemented to avoid, detect, address and prevent recurrence of serious noncompliance with GCP, the trial protocol and applicable regulatory requirements. ADS systems should be audited for their abilities to perform in this capacity and may need tuning to general (GCP) and specific (protocol, patient population, or intervention class) details.

Table 1: ICH GCP Principles and Considerations for ADS Use (Cont.)

Principles of ICH GCP (Number and Principal)	Considerations for ADS Use in Trials
9. Reliable Results	
9.2 Data to support good decision making	Data originating from ADS systems must maintain fidelity to the actual observation. ADS systems must undergo development, training, and testing to ensure accuracy and completeness in speech to text transcription and LLM interpretation and creation of trial documents and data points. The investigator and delegates should verify accuracy and completeness of these data and documents as required by the protocol.
9.3 Systems and processes for data capture, management, and analysis be fit for purpose	ADSs and surrounding processes – both system- and human-mediated - should be able to execute the complete and accurate capture of data required by the protocol. If the ADS cannot capture the data required, the investigator and qualified staff will capture the data using traditional means. It will be unlikely that all assessments and procedures are best captured via audio recording during an encounter, and there will most certainly still be data needing to be manually captured.
9.4 Robust record management for accurate reporting, interpretation, and verification	ADSs and surrounding processes – both system and human mediated – should ensure record integrity and traceability are maintained and that personal information is protected. Attribution – which party said which words and phrases and at which date and time should be captured by the ADS system and accurately and completely carried forth into clinical trial source documents and data systems.

Table 1: ICH GCP Principles and Considerations for ADS Use (Cont.)

Principles of ICH GCP (Number and Principal)	Considerations for ADS Use in Trials
9.5 Secure record retention and accessibility for inspection.	As discussed earlier, we postulate that the raw transcript or draft information contained in the ADS viewer and editor would not be considered an official record for the trial and is thus not subject to inspection or document retention timelines. Once this information is published or used to create a source document, earlier drafts should be locked for further editing and future reference should be given to the Source Documents contained outside of the ADS system. The Source Documents would then be considered essential records that need to be retained securely by sponsors and investigators for the required period in accordance with applicable regulatory requirements. These essential records should be available to regulatory authorities, monitors, auditors and IRBs/IECs (as appropriate) upon request to enable appropriate evaluation of the trial conduct to ensure the reliability of trial results.However, it is recommended that the information in the ADS system and the transcript be kept until the end of the trial and final database lock, just in case for reasons of data completeness and accuracy, the investigator would want to consult the original notes of the visit.
10. Clear and Documented Roles and Responsibilities	While the investigator may delegate visit duties to staff, he or she retains ultimate responsibility. ADS could be viewed as yet another delegate of responsibility contributing data for regulatory purposes and must be subjected to investigator oversight. The investigator is accountable for use of ADS to originate data for the protocol, for making sure this system is working properly and that he/she originates complete and accurate data required by the protocol.
11. Investigational product manufacture and handling (IP)	Use of the ADS should not supersede any other trial systems or processes related to IP receipt, storage, dispensing, and return. For instance, if an electronic randomization and trial supply management system (RTSM) is in use for the trial, use of the ADS should not alter or change the way it is used. If information discussed between the investigator/staff and about medication dispensing, use, accountability, or return is recorded and rendered in a visit summary, it may be consulted for purposes of updating source documents and other systems.

Artificial Intelligence Regulations

The following sections examine in depth the current AI regulations and regulatory opinions surrounding ADS in healthcare, in clinical research, and in investigational product development. Direct dialogue with regulators in individual geographies will be essential to obtain precise guidance on necessary regulatory oversight.

A topic of current and active debate, as well as emerging legislation, is whether ADS technologies should be regulated as Software as a Medical Device (SaMD). In the most definitive case, the National Health Service (NHS) of England issued a directive that any ADS used in its facilities be registered with the Medical Device with the UK Medicines and Healthcare products Regulatory Agency (MHRA) as a Class 1 medical device (47), and followed this up with a warning to its facilities indicating it will exercise enforcement (48).

In the United States, guidance is less direct; however, a recent FDA position paper on Gen-AI-enabled devices describes use cases and examples that fall into one of three groups (49):

1. **Not considered** a SaMD. Includes solutions "solely intended for the administrative support of a healthcare facility."
2. **May meet** the definition of a SaMD but FDA will use their discretion in enforcing. Includes "certain GenAI-enabled products that automate simple tasks for healthcare professionals."
3. **Unambiguously meets** the criteria of a SaMD for which the FDA will prioritize its oversight; discusses solutions with a direct role in clinical decision making that "could pose a risk to a patient's safety if the software were not to function as intended." The FDA paper continues and addresses ADS more directly:

"...hallucinations produced by a GenAI-enabled device can introduce uncertainty in the device's behavior, which can translate to difficulty in understanding the specific bounds of a device's intended use. For example, for a GenAI-enabled product that may be meant to summarize a patient's interaction with a healthcare professional, the possibility of that product hallucinating can present the difference between summarizing a healthcare professional's discussion with a patient and providing a new diagnosis that was not raised during the interaction."

A position paper by Dr. Hugh Harvey and Mike Pogose from the consulting firm Hardian Health objectively examines global regulations and accurately predicted the UK directive. This analysis suggests that Europe, where explicit guidance is still pending, may be stricter than in the UK and require registration as a Class II device (50).

The position of the Australian Therapeutic Goods Administration published a statement in August 2025, stating if ADS

simply transcribes a visit and takes notes of what was discussed, it is not considered a medical device; however, if it makes any analyses or interpretations that were not explicitly stated by the provider, it is considered a medical device (51). ADS used for clinical research could reasonably carry a "for research use only" designation.

Regulations for AI in Drug Development

In January 2025, the FDA released a draft guideline titled "Considerations for the Use of Artificial Intelligence to Support Regulatory Decision-Making for Drug and Biological Products" (52), outlining a seven step credibility assessment framework for AI models in drug development (53). This draft guidance document is available for public review and comment, and its content may be subject to change. However, it offers drug developers and clinical technology suppliers a valuable directional framework.

FDA's AI 7 Step Assessment Framework

1. **Question of Interest** answered by the model
2. **Context of Use** of the Model
3. **Level of Risk** when using the model
4. **Design a plan** to assess the model's credibility
5. **Conduct** the credibility assessment research
6. **Document the results** of the credibility assessment
7. **Submit the results** demonstrating the model's adequacy to the FDA

Proposed AI Validation Approach for ADS Used in Drug Development

We consider ADS technologies for the clinical trial use case within the FDA's proposed 7-step framework for AI used in drug development in their draft guideline to industry.

1 What is the **Question Of Interest** that the AI model addresses

The AI model determines **who** (which provider and which patient), **what** (the specific assessments), **when** (the time and date of the assessment), **how** (the method of assessment), and most importantly, what is **the result** or outcome of the assessment.

2 What is the **Context of Use** for the model?

The context of use is **a clinical trial, under a research protocol**, in a clinical or telehealth/home setting, in the conduct of an assessment or procedure within a particular disease or therapy area in a spoken language considering dialect and cultural manners of verbal expression.

Proposed AI Validation Approach for ADS Used in Drug Development

We consider ADS technologies for the clinical trial use case within the FDA's proposed 7-step framework for AI used in drug development in their draft guideline to industry.

3 What **Level of Risk** does the model carry?

Specifically, what is the level of influence (LOW or HIGH) on an outcome that the model has, and what is the magnitude of the consequence of that decision (LOW or HIGH)? Considering this use case, we assess that this model poses a "medium" level of risk. Figure 8 explains the logic.

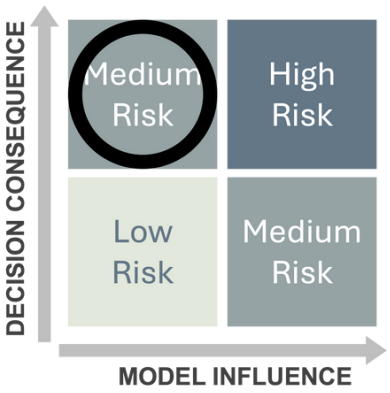


Figure 8. The fact that this model will originate clinical trial data gives it a HIGH decision consequence. The fact that the investigator has full review and editorial rights of all model outputs would give it a LOW model influence. Therefore, overall, use of ADSs could have a "medium" level of risk.

4 Designing the **Model Credibility Assessment Plan.**

Credibility assessment aims to determine if the model performs adequately and addresses the question of interest in the proposed COU. Credibility assessment should address the following aspects, and discussions with regulators should take place to ensure agreement with the approach and to confirm plans for documentation and submission for adequacy determination.

Proposed AI Validation Approach for ADS Used in Drug Development

4 Designing the **Model Credibility Assessment Plan.**

Elements to be addressed include:

Model Transparency	The operations of the model and all its components should demonstrate that they have been trained and tuned using data pertinent to the COU. This includes the types of data utilized for training, the representativeness of the data and the population from which it was derived, the sources of data, formats, and preprocessing steps, model architecture, training methodology, and ongoing validation, such as model updates. The transparency of how the outputs were produced should ensure the explainability and interpretability of the process.
Performance Evaluation in Real World Contexts	The testing and evaluation protocols should be thorough and provide strong evidence of performance meeting established success criteria. The model and human review output should be at least as good as what humans would produce without the model, and this must be demonstrated in real-world research settings. The metrics to measure and additional quality thresholds delineated in Figure 7 should be assessed.
Data Privacy and Security	Provide system design details of data security and encryption components, data from testing to confirm that it meets these standards, and details on how it will be fed into the model for ongoing training.
Human Oversight Processes	Consider activities occurring before, during, and after system use. Before using the system, individuals must establish an account with multi-factor authentication to ensure proper attribution and security. Processes must be implemented to guarantee GCP and regulatory compliance; only the PI or registered delegates are permitted to conduct visits for a given trial. Demonstrate performance of features that allow clinicians to review, edit, sign off, and receive alerts for discrepancies or issues needing resolution. Governance systems must be shown to ensure that the investigator's sign-off occurs before the information is used to create source data. Finally, it is crucial to demonstrate that using this system does not interfere with other obligations of the investigator or the quality of the investigation, and that its use is incorporated into quality systems, including those of the sponsor or CRO (for instance, the trial's monitoring plan).

Proposed AI Validation Approach for ADS Used in Drug Development

5 Conduct the Credibility Assessment

A qualified research team should perform credibility assessment studies under robust and detailed research protocols to generate the evidence discussed in step 4.

6 Document Credibility Assessment Results

The results of the credibility assessment tests should be documented in a Credibility Assessment Research Report.

7 Submit the Results to Regulators

ADS system owners are advised to submit their credibility assessment reports and obtain a certificate of approval from regulators. The sponsor may also wish to review these results and may even need to include the certificate of regulatory approval or even the reports in their market authorization application for the investigational product.

Credibility assessment should be an ongoing process. System performance should be monitored and measured regularly for accuracy and completeness, relevance and appropriateness, hallucinations, bias and fairness, and user experience and satisfaction. New features or enhancements should be evaluated for their impact on key performance indicators.

Inspection Readiness

The Appendix at the end of this document provides a detailed **Inspection Readiness Checklist** for the use of ADS technologies in FDA-regulated and GCP-compliant clinical trials.

This checklist considers the most conservative interpretation: the ADS system is regarded as an electronic source system, and the preliminary documents, such as the transcript and draft documents, are considered part of the source record.

Conclusion

Generative AI-enabled Ambient Digital Scribe technology is rapidly being adopted in healthcare to facilitate the passive documentation of interactions between physicians and patients, automate coding and billing processes, reduce the time and burden of charting, and enhance the overall quality of care.

This technology offers new ways to originate and acquire clinical trial data, enhance quality, detect adverse events, identify potentially eligible patients for trials, and other potential opportunities to enhance clinical research.

This paper aims to raise awareness of ADS in health systems that also operate as clinical research sites, and discuss the potential benefits, risks, opportunities, and regulatory considerations of ADS when used to generate clinical trial data, drug safety information, or patient insights.

This paper can also serve as a guide for research-exclusive institutions, which may not yet be familiar with these technologies and question their relevance for use in their institutions.

This paper can offer technology developers insights to develop new products or adapt existing products to meet clinical research needs.

The authors published this article as a white paper rather than a peer-reviewed journal article, as they desired to disseminate their thoughts on this topic more quickly than would be possible through a traditional peer-reviewed journal process.

The authors welcome comments and contributions and see this as the beginning of a fruitful collaboration to reduce the burden of clinical trials and the work-life balance of clinical research coordinators and investigators and develop novel ways to improve the quality and speed of clinical trial data generation.

Thank You For Reading!

Appendix: Inspection Readiness Checklist: Use of ADS Technologies in FDA-Regulated & GCP Compliant Clinical Trials

THE AUTHORS

Elise Felicione, MPH, MBA
CEO and Principal Consultant, Novatec G2
Founder, Ambient In Research
Web Profile

Craig Lipset, MPH
Founder, Decentralized Trial and Research Alliance
Co-Founder, Ambient In Research
CEO, Clinical Innovation Partners
Web Profile

Jeff Lee, MIM
Founder | Angel Investor
Web Profile

Jonathan Helfgott, MS
Johns Hopkins University, Senior Lecturer and Program Coordinator for MS in Regulatory Science & Founder of FDA Partners
Web Profile

Raj Ratwani, PhD, MPH
Vice President, Scientific Affairs, MedStar Health
Director, MedStar National Center for Human Factors Engineering in Healthcare
Professor, Georgetown University School of Medicine
Web Profile

Petros Okubagzi, MD
Vice President, Clinical and Translational Research
MedStar Health
Web Profile

Craig Serra, MS, MBA
Executive Data Management Leader, Amgen
Web Profile

Bert Hartog, PhD
Hartog Digital Health Consulting
Web Profile

Appendix:

Inspection Readiness Checklist: Use of ADS Technologies in FDA-Regulated & GCP Compliant Clinical Trials

System Validation & Regulatory Compliance (21 CFR Part 11; ICH E6(R2) §5.5.3)

□ The ADS system has been validated for its intended use in accordance with an established protocol demonstrating data integrity, security, traceability, reliability, and accuracy of transcriptions under study-relevant conditions.

□ The ADS system consistently produces accurate and complete records that are attributable to individual users.

□ The system maintains secure, computer-generated, time-stamped, non-overwritable audit trails that record the date and time of saved entries and any subsequent modifications.

□ Role-based access controls are implemented, and a current user access matrix is maintained and made available for the purposes of a site inspection.

□ All electronic records are retained in accordance with 21 CFR §312.62(c) and can be retrieved for inspection throughout the retention period.

□ If electronic signatures are used, the ADS system complies with 21 CFR Part 11 Subpart B, including linking signatures to their respective records.

Investigator Oversight and Source Documentation Responsibilities (21 CFR §312.60; ICH E6(R2) §4.1, §4.9)

□ PI maintains full oversight of the ADS-generated source documentation and confirms its accuracy and completeness.

□ PI or qualified designee reviews and finalizes each ADS originated document in a timely, documented manner, with a dated signature or equivalent certification.

□ A documented process exists for the PI to resolve discrepancies between ADS-generated transcripts and other clinical source data (e.g.-EHR).

Appendix:

Inspection Readiness Checklist: Use of ADS Technologies in FDA-Regulated & GCP Compliant Clinical Trials

□ The source data origin and control are clearly defined, including what version of the ADS originated document and data constitutes the official source.

□ All finalized source documentation is retained in a manner that ensures integrity, readability, and accessibility during monitoring and regulatory inspections.

□ Site staff using ADS systems have documented training that includes system use, roles, review processes, and Part 11 awareness.

Fit-for-Purpose Demonstration (FDA BIMO CPGM, ICH E6(R2) §5.0, §5.5; ICH E8(R1))

□ The sponsor or vendor has performed and documented a fit-for-purpose assessment of the ADS system for use in the specific clinical trial context.

□ The assessment includes metrics for transcription accuracy, completeness, usability in real-world clinical interactions, and limitations (e.g.- accents, ambient noise).

□ (In the case where the ADS system will be the source system of record) Performance data supports the reliability of ADS as a source for trial-critical data (e.g., AEs, medical history, concomitant medications).

□ The ADS system's role in the study is clearly described in the data flow diagram or equivalent system use document.

□ Monitoring and quality plans describe how ADS-generated source data are verified and integrated into the totality of clinical evidence.

□ A risk-based quality management approach has been used to evaluate and mitigate the impact of ADS-specific data risks (e.g., transcription errors).

Appendix:

Inspection Readiness Checklist: Use of ADS Technologies in FDA-Regulated & GCP Compliant Clinical Trials

Data Integrity, eSource Traceability, and ALCOA+ Compliance (FDA BIMO CPGM, ICH E6(R2) §2.10, §5.5.3; ICH E8(R1))

□ (In the case where the ADS system will be the source system of record), the ADS system supports ALCOA+ principles: data is attributable, legible, contemporaneous, original, accurate, complete, consistent, enduring, and available.

□ ADS transcripts are traceable from audio input to finalized document.

□ Any manual corrections or validations are documented and traceable to specific users.

□ The system architecture ensures data is not overwritten or lost and that historical saved data states can be reconstructed.

□ Procedures exist to manage discrepancies between ADS transcripts and other documentation or observations (e.g.-PROs).

Study-Level Planning and Documentation (ICH E6(R2) §5.5, §8.1; FDA BIMO CPGM 7348.811)

□ The use of ADS technology is explicitly described in study-level documents such as the protocol, DMP, or RBM plan(s).

□ ADS use is documented in the Trial Master File with version-controlled validation records, SOPs, and sponsor/vendor agreements. A reference should be made in the investigator site file to relevant institutional policies on ADS use.

□ The ADS system vendor provides technical documentation, including system architecture, data flows, change control procedures, and support models.

□ The PI and study site have been informed and have agreed to FDA's expectation of direct access to all ADS records, audit trails, and systems under 21 CFR §312.68.

□ Study monitors are trained and equipped to verify data captured via ADS and integrate it into SDV workflows.

References

(1) *Agency for Healthcare Research and Quality*. (2023, November). Physician burnout. Rockville, MD. https://www.ahrq.gov/prevention/clinician/ahrq-works/burnout/index.html

(2) West, C. P., Dyrbye, L. N., & Shanafelt, T. D. (2018). Physician burnout: Contributors, consequences, and solutions. *Journal of Internal Medicine*, 283(6), 516-529.

(3) *Centers for Disease Control and Prevention*. (2023, October). Health worker mental health. https://www.cdc.gov/vitalsigns/health-worker-mental-health/index.html

(4) National Center for Health Workforce Analysis. *Physician Workforce: Projections, 2022-2037*. (2024, November). https://data.hrsa.gov/topics/health-workforce/workforce-projections

(5) Porter, J., Boyd, C., Skandari, M. R., & Laiteerapong, N. (2023). Revisiting the time needed to provide adult primary care. *Journal of General Internal Medicine*, 38(1), 147-155.

(6) Khullar, D. (2023). Burnout, professionalism, and the quality of US health care. *JAMA Health Forum*, 4(3), e230024.

(7) Linzer, M., Manwell, L. B., Mundt, M., Williams, E., Maguire, A., McMurray, J., & Plane, M. B. (2005). Organizational climate, stress, and error in primary care: The MEMO study. In *Advances in Patient Safety* (Vol. 1)

8) Sinsky, C., Colligan, L., Li, L., & et al. (2016). Allocation of physician time in ambulatory practice: A time and motion study in 4 specialties. *Annals of Internal Medicine*, 165(11), 753–760.

(9) Butler, R., Monsalve, M., Thomas, G. W., & et al. (2018). Estimating time physicians and other health care workers spend with patients in an intensive care unit using a sensor network. *American Journal of Medicine*, 131(8), 972.e9–972.e15.

(10). Berg, S. (2024, August 13). Burnout on the way down, but “pajama time” stands still. *AMA*.

(11) *Elion Health*. AI ambient scribes products. Retrieved 06 June 2025, from https://elion.health/categories/ai-ambient-scribes/products

(12) Balloch, J., et al. (2024). Use of an ambient artificial intelligence tool to improve quality of clinical documentation. *Future Healthcare Journal*, 11(3), 100157.

(13) Biro, J. M., Handley, J. L., Mickler, J., Reddy, S., Kottamasu, V., Ratwani, R. M., & Cobb, N. K. (2025). The value of simulation testing for the evaluation of ambient digital scribes: A case report. *Journal of the American Medical Informatics Association*, 32(5), 928–931.

(14) Shah, S. J., et al. (2024). Ambient artificial intelligence scribes: Physician burnout and perspectives on usability and documentation burden. *Journal of the American Medical Informatics Association: JAMIA*, ocae295.

(15) Misurac, J., et al. (2024). The effect of ambient artificial intelligence notes on provider burnout. *Applied Clinical Informatics*, 10.1055/a-2461-4576.

(16) Tierney, A., et al. (2024). Ambient artificial intelligence scribes to alleviate the burden of clinical documentation. *NEJM Catalyst Innovations in Care Delivery*, 5(3), ocae295.

References

(17) Duggan, M. J., Gervase, J., Schoenbaum, A., et al. (2025). Clinician experiences with ambient scribe technology to assist with documentation burden and efficiency. *JAMA Network Open,* 8(2), e2460637.

(18). Shah, S., et al. (2025). Physician perspectives on ambient AI scribes. *JAMA Network Open*, 8(3), e251904.

(19) Leung, T. I., Coristine, A. J., & Benis, A. (2025). AI Scribes in Health Care: Balancing Transformative Potential With Responsible Integration. *JMIR Medical Informatics*, 13, e80898.

(20) Fierce Healthcare. (2024, December 18). 2025 outlook: What's next for AI scribes and virtual care. *Fierce Healthcare*. https://www.fiercehealthcare.com/ai-and-machine-learning/2025-outlook-whats-next-ai-scribes-and-virtual-care

(21) Landro, L. (2025, May 27). Why AI May Be Listening In on Your Next Doctor's Appointment. *The Wall Street Journal*. *https://www.wsj.com/health/healthcare/ai-ambient-listening-doctor-appointment-e7afd587*

(22) Trang, B. (2025, August 20). Epic's AI overhaul promises to address EHR headaches for clinicians and patients: Company introduces new AI assistants and ambient notes powered by Microsoft. STAT. Retrieved August 21, 2025, from STAT website: https://www.statnews.com/2025/08/20/epic-ehr-artificial-intelligence-microsoft/

(23) Wingo, P. (2025, January 23). AI clinician copilots: Technology, market trends, and key differentiators. *Elion Health*. https://elion.health/resources/ai-clinician-copilots

(24) Corneli, A., Pierre, C., Hinkley, T., Lin, L., Fordyce, C. B., Hamre, G., & Roe, M. T. (2017). One and done: Reasons principal investigators conduct only one FDA-regulated drug trial. *Contemporary Clinical Trials Communications*, 6, 31-38.

(25) Executive Steering Council & Partner Organizations of the Association of Clinical Research Professionals. (2025, March). White paper: Addressing the clinical research workforce crisis: A call for collective action.

(26) *Tufts Center for the Study of Drug Development*. (2021, January). Impact report.

(27) Bastek, A. (2021, September 15). Combating the burnout bottleneck among clinical research workers. *Stat*.

(28) Cagnazzo, C., Filippi, R., Zucchetti, G., & et al. (2021). Clinical research and burnout syndrome in Italy – only a physicians' affair? *Trials,* 22, 205.

(29) Biro, J., Handley, J. L., Cobb, N. K., Kottamasu, V., Collins, J., Krevat, S., & Ratwani, R. M. (2025). Accuracy and safety of AI-enabled scribe technology: Instrument validation study. *Journal of Medical Internet Research*, 27, e64993.

(30) Hose, B. Z., Handley, J. L., Biro, J., Reddy, S., Krevat, S., Hettinger, A. Z., & Ratwani, R. M. (2025). Development of a preliminary patient safety classification system for generative AI. *BMJ Quality & Safety*, 34(2), 130-132.

(31) Biro, J. M., Handley, J. L., McCurry, J. M., Visconti, A., Weinfeld, J., Trafton, J. G., & Ratwani, R. M. (2025). Opportunities and risks of artificial intelligence in patient portal messaging in primary care. *npj Digital Medicine,* 8(1), 222.

References

(32) Tran, B. D., Latif, K., Reynolds, T. L., Park, J., Elston Lafata, J., Tai-Seale, M., & Zheng, K. (2023). "Mm-hm," "Uh-uh": Are non-lexical conversational sounds deal breakers for the ambient clinical documentation technology? Journal of the American Medical Informatics Association, 30(4), 703-711.

(33) Duggan, M. J., Gervase, J., Schoenbaum, A., & et al. (2025). Clinician experiences with ambient scribe technology to assist with documentation burden and efficiency. JAMA Network Open, 8(2), e2460637.

(34) Rogers, W. A., Boltz, M., Mehta, S. S., Madigan, E. A., Kang, Y., Song, M., & Bowles, K. H. (2021). Characteristics of patients in randomized clinical trials compared with patients in a large health care system. JAMA Network Open, 4(4), e219310.

(35) National Academies of Sciences, Engineering, and Medicine. (2022). Improving representation in clinical trials and research: Building research equity for women and underrepresented groups. The National Academies Press.

(36) Wilkinson, M., et al. (2019). Baseline assessment of the evolving 2017 eClinical landscape. *Therapeutic Innovation & Regulatory Science*, 53(1), 71-80.

(37) Hamidi, M., et al. (2024). Source data verification (SDV) quality in clinical research: A scoping review. Journal of Clinical and Translational Science, 8(1), e101.

(38) De S. (2011). Hybrid approaches to clinical trial monitoring: Practical alternatives to 100% source data verification. Perspectives in Clinical Research, 2(3), 100-104.

(39) *Peterson Health Technology Institute*. (2025, March). Adoption of artificial intelligence in healthcare delivery systems: Early applications and impact: Report from the Peterson Health Technology Institute AI Taskforce.

(40) U.S. Food and Drug Administration. FDA record-keeping requirements for clinical investigators and sponsors developing drugs and biologics: 21 CFR 312.50, 312.58, 312.62, and 312.68.

(41). U.S. Food and Drug Administration. (2013 September). Electronic Source Data in Clinical Investigations, Guideline for Industry. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/electronic-source-data-clinical-investigations

(42) U.S. Food and Drug Administration. (2024 October). Electronic Systems, Electronic Records, and Electronic Signatures in Clinical Investigations: Questions and Answers. Guidance for Industry. https://www.fda.gov/media/166215/download

(43) Cohen, I. G., Ritzman, J., & Cahill, R. F. (2025). Ambient listening—Legal and ethical issues. *JAMA Network Open*, 8(2), e2460642.

(44) Parab, A. A., Mehta, P., Vattikola, A., & et al. (2020). Accelerating the adoption of eSource in clinical research: A Transcelerate point of view. *Therapeutic Innovation & Regulatory Science*, 54, 1141–1151.

(45) International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. (2025, January 6). *ICH harmonized guideline for good clinical practice E6, revision 3, final version*.

References

46) WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Participants. (2025, April 3). https://www.wma.net/policies-post/wma-declaration-of-helsinki/

(47) NHS England. (2025, April 29). Guidance on the use of AI-enabled ambient scribing products in health and care settings. https://www.england.nhs.uk/long-read/guidance-on-the-use-of-ai-enabled-ambient-scribing-products-in-health-and-care-settings/

(48) Moberly T. (2025, Jun 23). Doctors must stop using unregistered AI scribe tools, says NHS England. BMJ (Clinical research ed.), 389, r1302. https://doi.org/10.1136/bmj.r1302

(49) U.S. Food and Drug Administration. (2024, November 20). Total product lifecycle considerations for generative AI-enabled devices: *Executive summary for the Digital Health Advisory Committee*.

(50) Harvey, H., & Pogose, M. (2024, September 27). Are LLM-based ambient scribes and clinical summarizers medical devices? *Hardian Health Publications*.

(51) Therapeutic Goods Administration. (2025, August 14). Digital scribes (also referred to as AI scribes or ambient scribes). In How we regulate: Manufacture specific types of medical devices: Artificial Intelligence (AI) and medical device software. Australian Government Department of Health. Retrieved August 22, 2025, from https://www.tga.gov.au/how-we-regulate/manufacturing/manufacture-medical-device/manufacture-specific-types-medical-devices/artificial-intelligence-ai-and-medical-device-software/digital-scribes

(52) U.S. Food and Drug Administration. (2025, January). Considerations for the Use of Artificial Intelligence To Support Regulatory Decision-Making for Drug and Biological Products. Draft Guidance for Industry and Other Interested Parties. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-use-artificial-intelligence-support-regulatory-decision-making-drug-and-biological

(53) American Society of Mechanical Engineers. (2018). Assessing Credibility of Computational Modeling through Verification and Validation: Application to Medical Devices. ASME. V V 40.