Med Arcade’s White Paper

Version 2, July 2026 (supersedes the August 2025 edition)

Executive Summary

Healthcare sits on a $5 trillion data problem. Every clinical decision generates valuable data, yet fragmented records make it nearly impossible to capture, structure, or validate that information safely, driving delayed diagnoses, defensive medicine, and malpractice risk.

Consider a typical GP consultation: a patient with years of referral letters, specialist reports, and clinical notes scattered across fragmented systems. The doctor spends precious minutes, out of a typical 10 to 15 minute consultation, clicking through screens trying to piece together what's relevant. Critical information gets missed, decisions are made without full context, and when things go wrong there's no clear audit trail to show the decision was justified. This isn't just inefficient, it's dangerous.

Med Arcade unlocks that value by making clinical decisions safe, traceable, and defensible. We've built clinical decision support software for primary care, where 90% of patient interactions occur. Our platform integrates with legacy EHR systems and surfaces relevant patient information holistically in seconds. In our pilot at Cambridge Arbury Road Surgery, the BioSensum beta version reduced data review time by half and mouse clicks by fivefold, in some cases even more.

Getting this right means solving two hard problems. The first is extracting meaningful insight from years of fragmented medical history: what's clinically useful to one GP may not be to another, since different clinicians work differently, so getting it right takes many iterations and real-world feedback from working clinicians, not a one-off build. The second is building an interoperable, explainable AI platform that meets the regulatory and workflow requirements of primary care: passing regulatory review and building genuine interoperability both take real time and close engagement with practices.

Efficiency is just the entry point. While doctors use our platform, we capture high-fidelity clinical decision-making data showing not just what was decided, but why, the defensibility clinicians need and the validation infrastructure healthcare AI has so far lacked. The future of healthcare isn't just more AI, it's safe AI, built on validated clinical data. Med Arcade is making that future possible.

This white paper sets out the problem in more detail, how Med Arcade addresses it, its intended use in clinical practice, and the results from our pilot at Cambridge Arbury Road Surgery. It covers, in turn: the problem space, our product overview, intended use, and the Cambridge Arbury Road Surgery pilot.

Problem Space

General practitioners in England delivered a record 31.4 million appointments in June 2025 alone, almost a third more than before the pandemic (NHS England, 2025), yet every consultation carries hidden risk. The culprit isn't lack of data, it's that critical information is fragmented, buried, and impossible to trace when decisions go wrong.

The efficiency cost is stark: GPs make over 4,000 mouse clicks a day just to retrieve basic patient information, and spend 5 or more hours daily navigating fragmented records across multiple systems (Pinevich et al., 2021). Each GP is responsible for approximately 2,300 patients, making it nearly impossible to maintain comprehensive awareness of complex cases, especially as clinical workload per patient has risen sharply in recent years (de Dumast et al., 2024).

The cost isn't only time. When a clinical decision is later questioned, fragmented records make it hard to show the reasoning behind it, and the financial impact is real: clinical negligence claims cost the NHS £3.6 billion a year (2024 to 2025), up threefold since 2006 to 2007 (NHS Resolution, 2026).

Product Overview

Med Arcade is building BioSensum. It is a clinical intelligence software (MHRA-registered Class I Medical Device, GMDN: 61087) that integrates with electronic health record (EHR) systems to help doctors extract meaningful insights from complex patient records more quickly and accurately. We use an explainable AI architecture (developed in-house) to reduce hallucination risks and support data-driven medicine at the point of care.

Two hard problems we solve

Extracting meaningful insight from fragmented history: Years of medical history rarely resolve into a clean answer at the point of care. What's clinically meaningful to one GP may not be to another: different clinicians work differently, weigh information differently, and reach conclusions along different paths. Getting this right isn't a one-off engineering task; it takes many iterations and real-world feedback from working clinicians, which is why we've built BioSensum around continuous clinician input rather than a fixed specification.

Building for interoperability, safety, and explainability: An AI system that primary care can actually adopt has to be explainable to clinicians (showing data sources), interoperable with the EHR systems already in use, and able to pass NHS regulatory review. Each of these takes real time and cost. Regulatory review is a lengthy process by design, and genuine interoperability requires close ongoing engagement with the practices and systems we integrate with, not just technical connectors.

Dual value proposition

1. Immediate clinical efficiency: Transforms fragmented patient data across referrals, letters, labs, and historic notes into holistic visual summaries aligned with how GPs think and work. In our pre-integration pilot at Cambridge Arbury Road Surgery, we halved the time clinicians spent reviewing records, cutting the 4,000+ daily mouse clicks and 5+ hours spent navigating messy EHRs.

2. Safety infrastructure for healthcare AI: While delivering efficiency gains, our system captures high-fidelity clinical decision-making data showing which information informed each clinical decision. This creates the defensibility framework that protects clinicians from the £3.6 billion in annual negligence claims and reduces the 6 million harm events occurring in UK primary care. It also helps reduce missed diagnoses, estimated to cost the NHS £250 million annually (Cheraghi-Sohi et al., 2021).

Technical differentiation

BioSensum is not a wrapper or a generic large language model. It is built around a proprietary architecture purpose-built to work safely with fragmented longitudinal records, and refined through clinician feedback to fit primary-care workflows.

BioSensum combines several layers, each addressing a specific failure mode that generic AI systems still leave unsolved in regulated healthcare settings:

- **Ontology-grounded retrieval:** clinical queries are enriched using medical terminology and SNOMED coding, so the system retrieves genuinely relevant clinical evidence rather than a generic text match.
- **Staged processing for large patient histories:** very large or complex records are processed in stages, so clinicians receive a useful first-pass summary quickly, with the fuller record incorporated shortly after in the background.
- **Deterministic safety checks:** before any summary is generated, the retrieved evidence is checked against a fixed set of clinical safety rules, including allergy and medication conflicts, high-risk drug interactions, diagnosis conflicts, and timeline inconsistencies.
- **Evidence-validated generation:** outputs carry citations back to the source record. A validation step checks each citation against the underlying data and removes any that cannot be supported, so clinicians see claims that are traceable to original patient data.

This is a system-level combination of established and emerging techniques in retrieval-augmented generation, applied specifically to the safety and audit requirements of regulated healthcare environments (see Academic Foundations, below). It is this combination, not any single model, that is difficult to replicate without deep domain knowledge, clinical customisation, data access, and regulatory expertise. Early internal testing shows measurable improvements in retrieval precision and processing speed compared with generic retrieval methods, and we continue to test across a wider range of patient records and clinical scenarios.

The underlying intellectual property, our data-processing pipeline and clinical reasoning engine, is protected as proprietary software, with defensibility rooted in our technical architecture and its ongoing refinement against real clinical use.

Information flow – see Figure 1

- Input: the clinician provides a unique patient identifier via the user interface, optionally with the reason for the encounter (e.g. medication review), to help focus the summary.
- Data retrieval: patient-coded data is retrieved from the EHR system via dedicated APIs.
- Query enrichment: the encounter reason and clinical context are enriched using clinical terminology matching and SNOMED coding.
- Processing and retrieval: data is sanitised, processed, and formatted within BioSensum's secure UK-based cloud environment, structured into a hierarchical “virtual library”, and searched for the most clinically relevant evidence.
- Safety check: before generation, the retrieved evidence passes through a deterministic safety layer that checks for allergy-medication conflicts, high-risk interactions, diagnosis conflicts, and temporal inconsistencies.
- AI output: AI agents receive the validated evidence and generate a structured summary with inline citations.
- Evidence validation: each citation is checked against the source data, and any that cannot be supported are removed before display.
- Display: the validated, traceable summary is rendered in the user interface for the clinician to review.

Temporary data storage is strictly limited to runtime processing within secure UK-based cloud infrastructure.

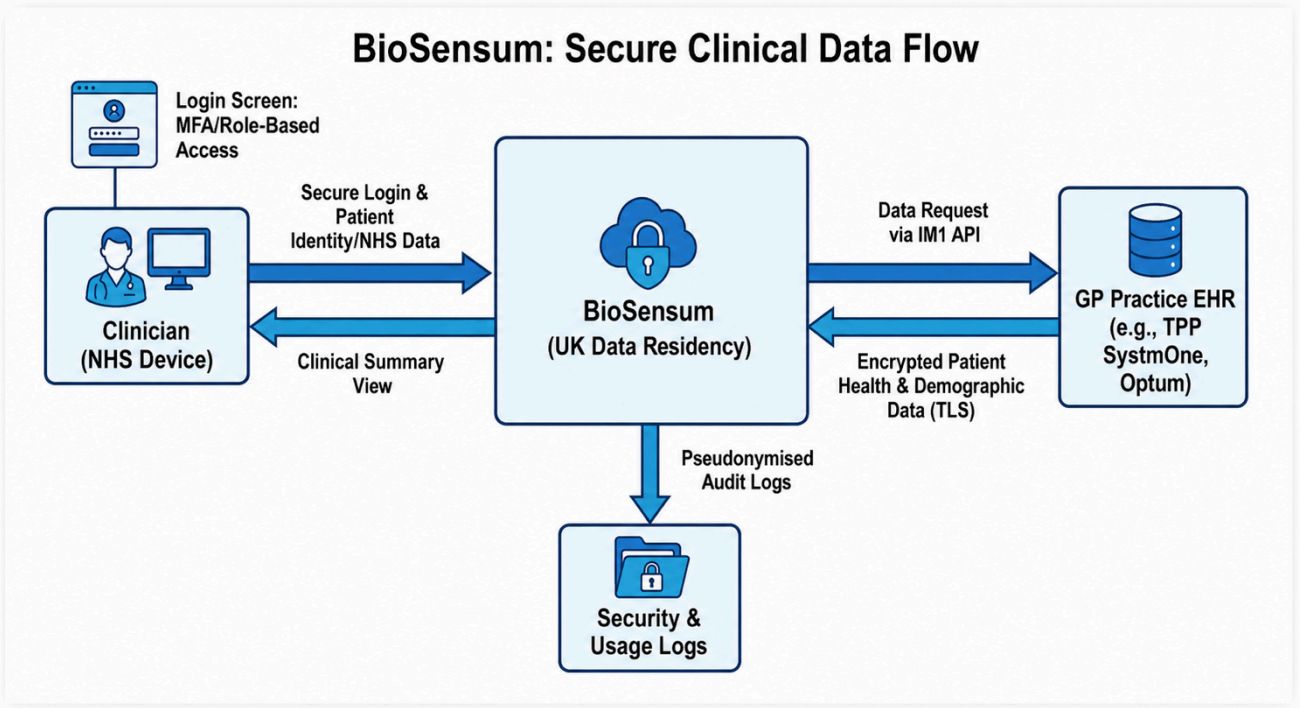

Figure 1: BioSensum secure clinical data flow, from clinician login through to pseudonymised audit logging. Patient data is processed in a secure UK-hosted environment with access controls, encryption, audit logging, and retention controls appropriate to the healthcare workflow.

IM1: Integration with NHS systems

BioSensum integrates with NHS systems through the NHS IM1 pathway, the standard compliance framework for third-party clinical software connecting to GP systems such as Optum and TPP SystmOne. We have completed the regulatory stack required to meet IM1 integration requirements, reflecting close alignment with NHS Digital's standards for clinical safety, data protection, interoperability, and technical assurance.

Intended Use

Use environment

Cloud-based web or desktop application used on GP practice computers alongside existing EHR systems (e.g. Optum, TPP SystmOne). Designed for GPs and specialist nurses during routine clinical activities: consultations, medication reviews, triage, and clinical assessments.

Example user journey

- Clinician logs into both the practice EHR and BioSensum on their GP practice computer.
- When reviewing a patient, they access the record using the unique patient identifier, optionally noting the encounter reason (e.g. medication review) to refine the summary.
- Software retrieves the patient's coded clinical data via NHS-approved APIs.
- Data is processed and sanitised on secure UK-based cloud infrastructure, checked for clinical safety conflicts, then submitted to BioSensum's AI model.
- The AI returns a structured, evidence-validated summary, a holistic history visible in seconds rather than minutes of clicking.
- Clinician reviews the summary during the patient encounter. The system logs what information was surfaced, creating an audit trail for defensibility.

- After consultation, the clinician proceeds to the next patient, with a traceable record of the clinical review maintained.

Fit with existing clinical activities

BioSensum complements existing clinical workflows without replacing clinical judgment. It integrates into routine use cases:

- Pre-consultation preparation: instant medical history summaries eliminate minutes of clicking through fragmented records.
- Real-time clinical assessments: holistic patient views during consultations, medication reviews, and triage.
- Investigation review: consolidated access to bloods, radiology, histopathology, and pathology results.
- Clinical correspondence processing: streamlined review of discharge letters and specialist clinic letters.
- Referral documentation: efficient drafting with comprehensive patient context readily accessible.

By transforming fragmented EHR data into structured, accessible formats, we reduce cognitive load, time per record, and the 4,000+ daily mouse clicks required to access clinically relevant information. Critically, the system ensures decisions are made with comprehensive context, whilst creating the audit trail needed to demonstrate clinical reasoning when outcomes are questioned. It does not automate or replace any clinical judgment or decision-making process.

Cambridge Arbury Road Surgery Pre-integration Pilot

This pilot, conducted in 2025, assessed whether the BioSensum beta version improves GP efficiency and decision-making when reviewing EHRs. The core aim was to compare performance with and without the tool, using anonymised patient records.

This was an early, small-sample pilot (4 GPs, 8 patient records). We treat it as directional evidence rather than a definitive validation, and a larger pilot will be carried out for further validation as we scale deployment.

Methods

Four GPs took part in this study, each using the same computer setup in a controlled environment. Eight anonymised, consented patient records were used to simulate realistic primary care scenarios, with clinical questions and gold-standard answers for each, predefined by Dr Pheng Toh, the GP who observed the study. These records formed the basis for evaluating performance with and without the Med Arcade solution.

1. Baseline task: to control for individual differences in baseline review speed and style, each GP first reviewed 2 neutral “practice” records without Med Arcade. These records were excluded from the main analysis; their performance, especially time per task and click count, was used to normalise final results.

2. Patient record assignment: each GP then reviewed 8 experimental records, 4 with Med Arcade and 4 without. Records were randomly assigned but counterbalanced across GPs to avoid learning or fatigue bias (Figure 2).

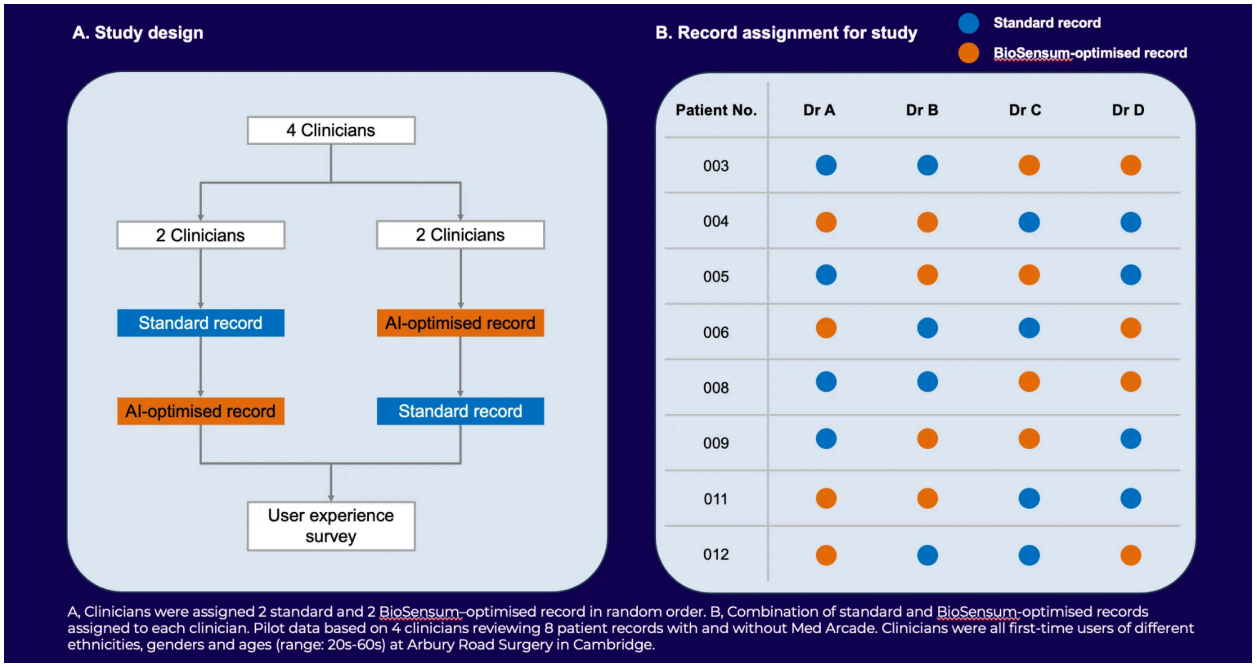

Figure 2: Study design and record assignment. Clinicians were first-time users of different ethnicities, genders, and ages (range 20s to 60s) at Cambridge Arbury Road Surgery.

3. Review tasks: GPs answered standardised questions per record, for example "How many tonsillitis episodes occurred in the last 12 months?", "What antibiotics has the patient had in the last 6 months?", and "What is the current care plan?" Each task was timed automatically, mouse clicks were recorded using browser logging or screen-tracking software, and accuracy was measured against predefined answers.

4. Outcomes measured: primary outcomes were time to complete task, exploration of records (click count), and accuracy score. Secondary outcomes were GP confidence and perceived usefulness, via a short survey.

Controls: GPs were unfamiliar with the records beforehand; task complexity was kept consistent across records; identical computer setups were used to avoid bias; and the timing and logging system was piloted before the study started.

Results

The BioSensum beta version reduced patient case review time by up to 71% and mouse clicks by up to 88% (Figure 3).

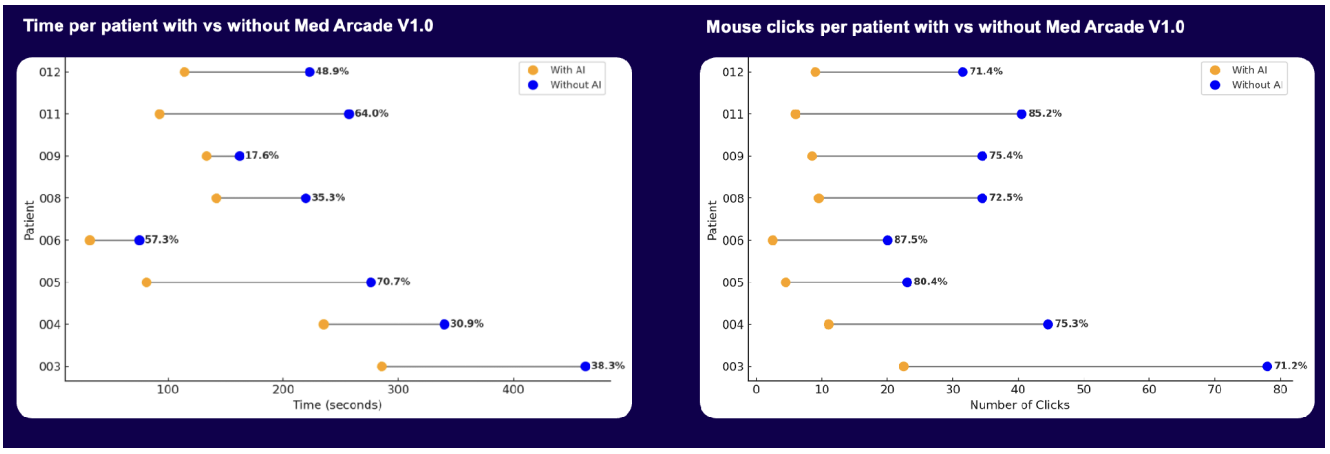

Figure 3: Time and mouse clicks per patient, with vs without the BioSensum beta version, across all 8 patients in the pilot.

Value of the solution was even more apparent in more complex cases (i.e. those with a longer standard review time): time saved correlated with standard review time at r = 0.72 (p = 0.043), and clicks saved correlated at r = 0.84 (p = 0.009) (Figure 4).

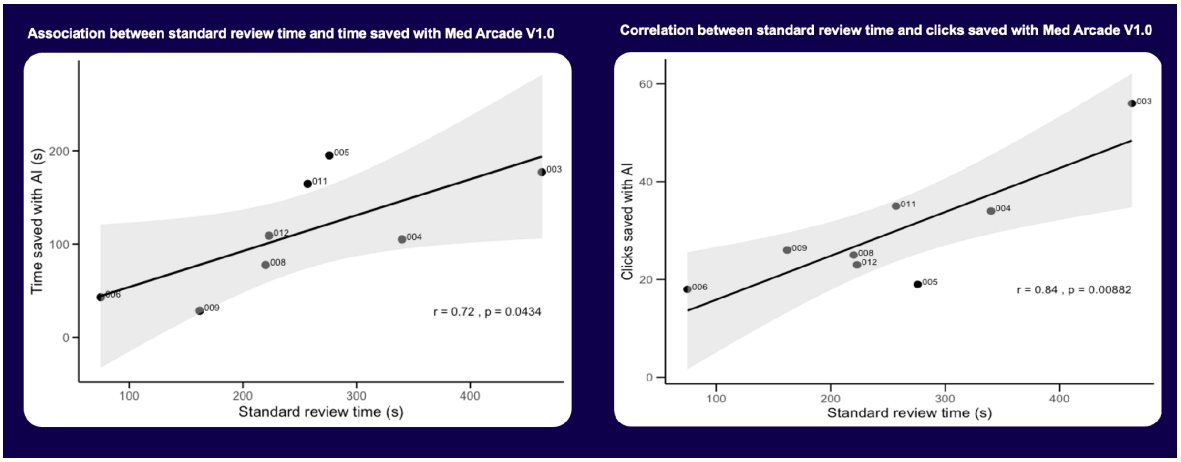

Figure 4: Association between standard review time and time/clicks saved with the BioSensum beta version.

Other key findings

In a case illustrating the potential of BioSensum to enhance patient safety, a patient with multiple attendances across A&E and rheumatology clinics underwent several investigations recorded in disparate sections of the health record. Clinicians using only the legacy system did not identify an abnormal test result, despite its inclusion in a clinic letter. In contrast, clinicians who used BioSensum alongside the legacy system successfully detected the abnormality.

100% of clinicians who tested the product asked for the paid version. Following the pilot, we've built a waitlist of over 10 GP practices for early access.

This market need has also been independently validated: an innovation review panel convened by an NHS health innovation network confirmed the clinical need for BioSensum and recommended we prioritise health economic validation and pilot deployment with interested partners, which is the direction our next phase of work is taking.

Examples of other data points we collected on clinician behaviour (not exhaustive):

- Information-seeking patterns: which sections of the record clinicians access first and most frequently.
- Search behaviour, both for key words and files, during record review.
- Drafting of referral letters.
- Responses to clinical questions.
- Questions in mind for individual patient cases.

Academic Foundations

Our approach builds on established and emerging research in retrieval-augmented generation and clinical AI, rather than on any single proprietary model. Relevant published work includes:

- Lewis et al., Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks, NeurIPS, 2020.
- Gao et al., Retrieval-Augmented Generation for Large Language Models: A Survey, 2024.
- Asai et al., Self-RAG: Learning to Retrieve, Generate, and Critique through Self-Reflection, 2024.

- Abo El-Enen et al., A Survey on Retrieval-Augmented Generation Models for Healthcare Applications, 2025.
- Guan, Medical Terminology Definition-Enhanced RAG for Hallucination Mitigation, 2025.

Med Arcade's contribution is the system-level combination of these techniques into a single, safety-checked pipeline for longitudinal primary care records, rather than any one algorithm in isolation.