FDA Clears First LLM as a Medical Device: Inside UpDoc's 510(k)

Published July 3, 2026 35 min read


FDA Clears First LLM as a Medical Device: Inside UpDoc's 510(k)

Executive Summary

On December 23, 2025, the U.S. Food and Drug Administration (FDA) cleared **UpDoc, Inc.**'s type 2 diabetes management software under submission number **K253281**, a decision the agency's own database shows was filed on September 29, 2025 (Source: innolitics.com). UpDoc did not publicize the clearance until June 25, 2026, when it announced **$18 million** in oversubscribed seed financing from backers including the American Diabetes Association, Eli Lilly and Company, and Mayo Clinic, alongside pilot deployments at Cleveland Clinic, Allegheny Health Network, and UCSF Health (Source: prnewswire.com). The company describes UpDoc as the first FDA clearance for **software as a medical device (SaMD)** that uses **patient-facing large language models (LLMs)** (Source: prnewswire.com). The underlying FDA decision summary, however, describes something narrower: a prescription-only tool that logs blood glucose, meal, symptom, and adherence data through a **Conversation Service (UpDoc Agent)** and executes insulin-dosing instructions defined entirely by a physician, not the model itself (Source: accessdata.fda.gov).

This report examines what FDA actually authorized, how UpDoc's regulatory team engineered a **510(k)** clearance around a technology many assumed would require the slower, costlier **De Novo** or **Premarket Approval (PMA)** pathway, and what the decision means for the roughly 221 companies now competing in the AI medical device space each year (Source: innolitics.com). The clearance rests on **substantial equivalence** to Hygieia's **d-Nav System** (K181916), a non-conversational insulin dose calculator cleared in 2019 under product code NDC, "Calculator, Drug Dose," classified within **21 CFR 868.1890** (Source: accessdata.fda.gov) (Source: innolitics.com). Because the predicate's intended use, insulin-dose recommendation for adults with type 2 diabetes, was preserved and the conversational layer was treated as a data-capture and communication interface sitting outside a locked, deterministic dosing algorithm, FDA did not require a new device classification (Source: innolitics.com). UpDoc also secured a **Predetermined Change Control Plan (PCCP)**, the mechanism FDA finalized in December 2024 that lets manufacturers pre-authorize future algorithm and interface updates without new marketing submissions, provided modifications preserve "deterministic insulin dosing logic without altering core clinical decision-making" (Source: accessdata.fda.gov) (Source: mcdermottplus.com).

The clearance lands inside a regulatory system already processing AI at industrial scale. FDA had authorized **1,451** AI- or machine-learning-enabled devices by the end of 2025, up from just **6** in 2015, with a record **295** clearances in 2025 alone; Innolitics separately reports that the median time to clearance across that 2025 cohort was **142 days**, with an average of 150 days (Source: theimagingwire.com) (Source: innolitics.com). Roughly 97% of these clearances still travel through 510(k), and independent researchers have found persistent gaps in premarket evidence: a JAMA Health Forum study of 691 FDA-cleared AI/ML devices found only **1.6%** cited randomized clinical trial data and under **1%** reported patient outcomes (Source: pmc.ncbi.nlm.nih.gov). UpDoc's own decision summary lists no clinical study data, marking usability and software verification, not clinical efficacy, as the evidentiary basis for clearance (Source: accessdata.fda.gov).

Reaction has split sharply along the line separating what FDA cleared from what UpDoc marketed. Cleveland Clinic's chief digital officer, Rohit Chandra, said the clearance "was an important factor" in the decision to pilot the platform (Source: linkedin.com), while regulatory consultant Yujan Shrestha called the clearance a "cage built around" the LLM rather than authorization of the model itself (Source: innolitics.com). STAT News headlined its coverage of the episode with the question "Is the LLM an interface or the decision-maker?" (Source: statnews.com). This report walks through the mechanics of the clearance, the PCCP that governs its evolution, comparable milestones such as Digital Diagnostics' 2018 autonomous De Novo clearance and FDA's still-unresolved deliberations on generative AI mental health devices, and what the precedent means for developers pursuing FDA clearance for LLM-enabled clinical tools in 2026 and beyond.

Introduction and Background

For most of the past decade, "**AI medical device**" has meant a radiology algorithm that flags a nodule or an ECG classifier that scores arrhythmia risk. As of the end of 2025, radiology still accounted for **76%** of the 1,451 cumulative FDA authorizations on the agency's AI/ML-enabled device list, with cardiovascular software a distant second at roughly **9%** (Source: theimagingwire.com) (Source: intuitionlabs.ai). Almost none of that population talked back to a patient. UpDoc's clearance changes that picture, at least at the margins: the FDA's own **AI-Enabled Medical Device List** page states the agency will "explore methods to identify and tag medical devices that incorporate foundation models encompassing a wide range of AI systems, from large language models (LLMs) to multimodal architectures," and specifically invites sponsors to flag **LLM**-based functionality in future submissions (Source: fda.gov). UpDoc's June 2026 announcement was timed to claim that distinction publicly.

The company itself is not new. UpDoc went public with its concept in **January 2024**, describing a clinician-directed conversational AI platform for medication management built on multiple large language models, including GPT-4 through Azure OpenAI and Google Cloud's MedLM and Vertex AI models, according to contemporaneous trade reporting (Source: innolitics.com).The clinical root of the product traces to **MIVA**, a Stanford-led randomized trial registered as **NCT05081011**, which studied voice-based conversational AI for basal insulin titration and published results in JAMA Network Open in December 2023 (Source: innolitics.com). UpDoc's founders, CEO Sharif Vakili and Chief Technology Officer Ashwin Nayak, led that trial at Stanford Medicine before founding the company, filed a cluster of conversational-medication-management patents through late 2025, and then pursued FDA clearance quietly: the 510(k) was received on September 29, 2025 and cleared on December 23, 2025, months before the company said anything publicly (Source: medicaldesignandoutsourcing.com) (Source: innolitics.com).

That gap between a quiet regulatory clearance and a loud commercial launch is itself instructive for anyone trying to answer the question "**fda clears first llm as a medical device**": what got marketed as a category-defining breakthrough in "agentic AI" was, in the FDA's own paperwork, an incremental modification to an established, narrowly defined device type, a **drug dose calculator**. This report reconstructs the clearance from FDA's public 510(k) decision summary, cross-checks it against contemporaneous legal, trade, and regulatory analysis, and situates it inside a broader FDA framework, spanning **510(k)** substantial-equivalence review, **De Novo** classification, **PMA**, and the newer **PCCP** mechanism, that is already processing hundreds of AI-enabled clearances a year (Source: fda.gov). The goal is to give clinical AI developers, health system counsel, and life-sciences strategy teams a precise, source-verified account of what changed, what did not, and what the precedent implies for the next wave of generative AI submissions.

What FDA Actually Cleared: Inside the UpDoc 510(k)

The Device, the Predicate, and the Product Code

K253281's **Indications for Use** statement is unambiguous about scope: "UpDoc is a software as a medical device (SaMD) intended to provide medication management for patients aged 18 years and older who have been diagnosed with type 2 diabetes," delivering "insulin treatment plan instructions based on a healthcare provider-specified treatment plan" (Source: accessdata.fda.gov). The device is explicitly contraindicated for patients with type 1 diabetes and is restricted to prescription use, "Rx - For Prescription Use Only" (Source: accessdata.fda.gov). Architecturally, the decision summary describes three modular components: a provider-facing web portal for configuring dosing instructions and safety protocols, a patient mobile application for logging data via manual entry, voice, or text, and a cloud-based back end split into a "**Conversation Service (UpDoc Agent)**" and a separate "**Clinical Service**" that computes insulin instructions from provider-defined parameters (Source: accessdata.fda.gov). Notably, the manufacturer states in its own filing that "UpDoc does not interpret or diagnose symptoms," and that any symptom outside a pre-defined treatment protocol triggers a system lock instructing the patient to seek medical attention rather than a model-generated clinical judgment (Source: medicaldesignandoutsourcing.com).

That separation is the regulatory hinge of the entire clearance. UpDoc's predicate, Hygieia's **d-Nav System** (K181916, cleared February 4, 2019), is a non-conversational insulin dose calculator with the same product code, NDC ("Calculator, Drug Dose"), and the same classification regulation, 21 CFR 868.1890 ("Predictive pulmonary-function value calculator," a regulation whose name predates its current use for dosing software) (Source: accessdata.fda.gov). Independent regulatory analysis from device consultancy Innolitics found that "UpDoc's 510(k) summary says the two devices share the same intended use: software systems that determine the next insulin dose recommendation to aid insulin management," and concluded that "the addition of the conversational agent and other generative AI components did not trigger a De Novo" precisely because the dosing logic itself, not the conversational shell around it, defines the regulated function (Source: innolitics.com) (Source: innolitics.com). Notably, d-Nav itself was cleared without an authorized PCCP, underscoring how much additional regulatory machinery UpDoc layered onto essentially the same clinical claim six years later, and Hygieia, the predicate's developer, later filed for bankruptcy in 2024, before UpDoc's clearance was even finalized (Source: accessdata.fda.gov) (Source: medicaldesignandoutsourcing.com).

The decision summary is also notably thin on clinical evidence. Under "Performance Characteristics," FDA lists non-clinical performance as "Not applicable" and clinical studies as "Not applicable," relying instead on human factors and usability engineering data and software documentation aligned to a "major level of concern" given the device's role in insulin dosing (Source: accessdata.fda.gov) (Source: accessdata.fda.gov). Reviewers cited **IEC 62304** and FDA's guidances on software content, cybersecurity, and human factors engineering as the applicable standards framework (Source: accessdata.fda.gov). STAT News summarized the regulatory posture bluntly: the device "is regulated in the same product category with drug dose calculators that take inputs like blood glucose levels and return insulin dosing recommendations," with the LLM layer functioning as an intake and communication interface rather than an independent clinical decision-maker (Source: statnews.com).

The Predetermined Change Control Plan

A second, equally important element of the clearance is procedural rather than clinical: UpDoc's 510(k) filing explicitly states its purpose includes to "Establish a Predetermined Change Control Plan (PCCP)," alongside the new-device submission itself (Source: accessdata.fda.gov). A PCCP is the mechanism FDA finalized in final guidance on December 3, 2024, allowing manufacturers of AI-enabled devices to pre-specify future modifications, and the associated validation methodology, so that later updates can ship without triggering a fresh marketing submission (Source: mcdermottplus.com). The guidance requires at least three components: "a description of modifications, a modification protocol, and an impact assessment" (Source: mcdermottplus.com).

UpDoc's decision summary states that "the sponsor proposes the following six categories of modifications in their PCCP plan," though the released document enumerates five: adjustments to default clinical values (such as fasting glucose targets or maximum configurable insulin doses), updates to insulin product references as new formulations or biosimilars reach market, addition of new insulin-dosing features such as carb counting, user interface and language localization changes, and support for alternative data input methods such as Apple HealthKit or new continuous glucose monitor integrations (Source: accessdata.fda.gov). Crucially, the plan's guardrails explicitly wall off the clinical reasoning itself: performance requirements "mandate that all modifications maintain deterministic insulin dosing logic without altering core clinical decision-making," with "zero tolerance for deviation" in data import accuracy (Source: accessdata.fda.gov). In other words, UpDoc can iterate its conversational interface and expand its supported medication list under the PCCP; it cannot use the PCCP to let the underlying LLM make freer dosing judgments without a new submission. Industry-wide, PCCPs remain a minority practice even among AI-cleared devices: Innolitics' analysis of 2025 clearances found that "in 2025, **30 devices (10.2%)** were cleared with a PCCP," making UpDoc's inclusion of one a deliberate, above-average investment in lifecycle regulatory planning (Source: innolitics.com).

Key Changes: How This Clearance Redraws the AI Regulatory Playbook

From Machine Learning Classifiers to Patient-Facing LLMs

FDA's AI/ML regulatory apparatus was built for a different kind of algorithm than a conversational LLM. The agency's foundational **AI/ML SaMD Action Plan**, published in January 2021, and its earlier 2019 discussion paper on modifications to AI/ML-based SaMD, were designed around locked or incrementally learning classifiers that output a score or a flag, not a system generating open-ended natural-language responses to patients (Source: fda.gov). Since then FDA has layered on additional guidance largely aimed at machine-learning models: **Good Machine Learning Practice (GMLP)** guiding principles in October 2021, draft PCCP guidance in April 2023, PCCP guiding principles in October 2023, transparency guiding principles for ML-enabled devices in June 2024, and finally the PCCP final guidance in December 2024 (Source: fda.gov). None of that stack was written with generative, patient-conversant LLMs specifically in mind, which is why FDA's January 6, 2025 draft guidance, "**Artificial Intelligence-Enabled Device Software Functions: Lifecycle Management and Marketing Submission Recommendations**," explicitly asked for public comment on "the adequacy of the recommendations to address concerns that may be raised by emerging technology such as generative AI" (Source: fda.gov). At the time that draft was issued, Troy Tazbaz, director of FDA's Digital Health Center of Excellence, noted that "the FDA has authorized more than 1,000 AI-enabled devices through established premarket pathways," underscoring how mature the non-generative side of the pipeline already was even as the generative side remained an open regulatory question (Source: fda.gov).

UpDoc's clearance did not wait for that draft guidance to finalize. Instead, its regulatory strategy worked entirely within the existing 510(k) framework by treating the LLM as a data-capture and communication layer sitting outside the regulated dosing function. Innolitics summarized the pattern this way: "the strongest public evidence still points to clinical decisions operating inside healthcare-provider-set limits, with the LLM layer wrapped around a narrower regulated function... this is likely the reason FDA allowed this technological jump as a 510(k) rather than a De Novo" (Source: innolitics.com). The same analysis concluded that "this is a pretty narrow and well-risk-controlled use case as far as agentic AI goes," a characterization sharply at odds with UpDoc's own marketing language describing "physician-grade agentic AI" and clinical AI agents that "complete tasks autonomously" (Source: innolitics.com) (Source: prnewswire.com).

The 510(k) Predicate Strategy That Made It Possible

The core regulatory lesson of K253281 is a predicate strategy, not a new statutory category. A **510(k)** clearance requires demonstrating that a new device has "the same intended use" as a predicate and either the same technological characteristics or different characteristics that "do not raise different questions of safety and effectiveness," with FDA typically issuing a substantial equivalence determination within about 90 days of a complete submission (Source: fda.gov) (Source: fda.gov). By anchoring UpDoc's intended use to d-Nav's narrow, already-cleared claim, insulin dose

recommendation for type 2 diabetes, and by keeping the conversational agent's outputs confined to data collection, formatting, and delivery of provider-configured instructions, UpDoc's team avoided the harder argument that a De Novo or PMA filing would require: proving from scratch that an LLM-driven system is safe and effective as an independent clinical decision-maker.

This is consistent with a well-documented industry pattern. A JAMA Health Forum analysis of the broader AI/ML device population found that "one-third of AI/ML devices cleared through the 510(k) pathway originate from non-AI/ML devices," meaning predicate-based incremental claims, not novel clinical validation, are already how most algorithmic functionality enters the market, and the same analysis found that its 691-device sample "were cleared through the 510(k) (668 [96.7%]), De Novo (n = 20 [2.9%]), or PMA (n = 3 [<1%]) pathways" (Source: pmc.ncbi.nlm.nih.gov) (Source: pmc.ncbi.nlm.nih.gov). The same dynamic that let a CT-image classifier claim equivalence to a non-AI radiology workstation a decade ago let UpDoc claim equivalence to a non-conversational insulin calculator in 2025. Innolitics' post-clearance "Builder Playbook" distills the strategy into concrete guidance for other developers: "pick one narrow regulated function before designing the agent," "identify the predicate, classification, and special controls early," "convert free-form conversation into structured fields with schema checks, confirmations, and audit logs," and "put deterministic clinical logic behind the structured-data boundary" unless a sponsor is prepared to pursue a De Novo submission instead (Source: innolitics.com) (Source: innolitics.com).

The broader product-code family UpDoc used is more elastic than its 1970s-era name, "Predictive pulmonary-function value calculator," suggests. Innolitics notes that "868.1890 already contains several calculator-style software product codes beyond classic pulmonary-function calculators," including PHY for sparse-sample pharmacokinetic dosing software and PDT for burn-resuscitation decision-support software, concluding that "the category can accommodate software that calculates treatment-relevant outputs from structured inputs, especially where the output is bounded by an algorithm, protocol, model, or clinician-configured plan" (Source: innolitics.com). For developers researching "**drug dose calculator 510k predicate device**" strategies, this product-code family, not a bespoke AI classification, is the practical starting point for medication-management SaMD with conversational front ends.

Regulatory Guardrails: PCCP, Human-in-the-Loop, and State Practice-of-Medicine Law

Legal analysis of the clearance converges on one theme: the human clinician, not the software, remains the accountable decision-maker. McGuireWoods' health regulatory practice concluded that "clinical AI will likely continue to serve as a support tool for clinicians due to 'human-in-the-loop' and state law practice-of-medicine requirements, rather than be approved to make independent diagnostic or treatment decisions" (Source: mcguirewoods.com). The firm's alert frames K253281 as establishing "a viable regulatory pathway for SaMD products incorporating patient-facing LLMs, though future approvals will depend on each product's intended functions, safety and efficacy," and cautions that developers "should also monitor FDA's evolving expectations for AI-enabled devices, including post-market surveillance, transparent labeling and risk communication, cybersecurity, data integrity and other controls, and predetermined change control plans" (Source: mcguirewoods.com) (Source: mcguirewoods.com).

That caution is echoed in FDA's own internal deliberations elsewhere in the agency. In November 2025, before UpDoc's clearance became public, FDA's Digital Health Advisory Committee (DHAC) convened specifically to weigh the risks of generative AI mental health devices, considering a hypothetical "therapy device" built on an LLM "with unique outputs that mimic a conversation with a human therapist" (Source: statnews.com). A summary of that meeting from Hyman, Phelps & McNamara noted that as of December 2025, "FDA has authorized more than 1,200 artificial intelligence (AI)-based digital devices for marketing. To date, none of these has been indicated to address mental health," and that the committee flagged specific generative-AI risks, "hallucination" and "sycophancy," that do not arise in classical ML classifiers (Source: thefdalawblog.com) (Source: thefdalawblog.com). FDA speakers at that meeting "proposed the inclusion of a predetermined change control plan (PCCP) and a performance monitoring plan in premarket submissions as potential strategies to mitigate risks arising from changes in AI-enabled software performance over time," the same mechanism UpDoc had already secured weeks earlier for a much narrower use case (Source: thefdalawblog.com). The juxtaposition is instructive for anyone researching "**fda regulation of large language models in healthcare**": FDA was actively debating generative AI's mental-health risk profile in the same season it cleared a bounded, provider-governed generative AI tool for diabetes management.

Implementation Considerations and Process Changes for AI Medical Device Developers

Developers asking "**how to get fda clearance for an ai medical device**" that incorporates conversational or generative components can draw a fairly specific playbook from the UpDoc precedent and the surrounding guidance ecosystem, though every element below should be validated against a sponsor's specific intended use through FDA's Q-Submission (pre-submission) process rather than assumed to transfer automatically:

- **Anchor the intended use to an existing, narrow predicate.** UpDoc's clearance worked because its indications for use mirrored d-Nav's nearly word for word; the conversational agent did not expand what the device claimed to do, only how data reached it (Source: accessdata.fda.gov).

- **Keep the LLM outside final clinical decision authority.** FDA's decision summary repeatedly emphasizes that insulin instructions are "computed... based on the healthcare provider-defined treatment parameters," not generated by the conversational layer (Source: accessdata.fda.gov).
- **Build a PCCP before, not after, launch.** Only **10.2%** of 2025 AI/ML clearances included one, meaning most competitors will lack this flexibility and will face a new submission for each significant update (Source: innolitics.com).
- **Invest early in human factors and usability testing.** With no clinical study in the file, UpDoc's clearance leaned on "comprehensive human factors engineering protocols and study results demonstrating that intended users can safely and effectively perform all critical tasks" (Source: accessdata.fda.gov).
- **Treat prompts, models, and thresholds as controlled software artifacts** subject to the same verification, validation, and change-control discipline as any other "major level of concern" software function (Source: accessdata.fda.gov).
- **Budget for a multi-quarter, not multi-week, review.** UpDoc's own submission took roughly 12 weeks from receipt to clearance (September 29 to December 23, 2025), while some 2025 AI/ML clearances industry-wide "took over 200 days, highlighting the complexities and challenges that can arise with more novel or complex technologies" (Source: innolitics.com) (Source: innolitics.com).
- **Plan for parallel compliance obligations beyond FDA.** McGuireWoods flags HIPAA, data governance, consent, cybersecurity, and vendor-contract liability as separate, non-optional workstreams for any LLM system with access to patient data (Source: mcguirewoods.com).

Sponsors evaluating whether their own product fits a **510(k)** predicate strategy, a **De Novo** classification request for products with no suitable predicate, or a full **PMA** for higher-risk claims should treat that pathway decision as the single highest-leverage regulatory choice in the product roadmap; FDA's own guidance stresses that "many changes to artificial intelligence and machine learning-driven devices may need a premarket review" precisely because the agency's traditional device paradigm "was not designed for adaptive artificial intelligence and machine learning technologies" (Source: fda.gov). Life-sciences consultancies advising clinical AI developers on this decision, including firms such as **IntuitionLabs.ai** that track FDA's device authorization data at scale, generally counsel sponsors to treat the predicate search, product-code selection, and PCCP scope as sequential, interdependent decisions made before software architecture is finalized, rather than retrofitted onto a nearly complete product, since each choice narrows or widens the evidentiary burden that follows (Source: intuitionlabs.ai).

Data Analysis and Evidence

UpDoc's clearance sits inside an FDA AI device pipeline that has grown at a pace few other medical technology categories can match. The agency authorized just **6** AI/ML-enabled devices in 2015; that figure rose to **91** in 2022, **221** in 2023, **253** in 2024, and a record **295** in 2025, pushing the cumulative authorized total to **1,451** by year-end 2025 (Source: intuitionlabs.ai) (Source: theimagingwire.com). An earlier snapshot from MedTech Dive, using data current to August 7, 2024, put the cumulative count at **950** devices and separately documented that "About 97% of AI-enabled devices on the list were 510(k) cleared as of August 2024," with only 22 devices via De Novo and just 4 via PMA in the entire history of the list at that point (Source: medtechdive.com) (Source: medtechdive.com). *Table 1* below traces that growth trajectory and shows how consistently the 510(k) pathway, not De Novo or PMA, has carried the category forward, a pattern UpDoc's own clearance simply extended into conversational AI.

Table 1: FDA AI/ML-Enabled Medical Device Authorizations, Selected Years

YEAR	NEW AUTHORIZATIONS	CUMULATIVE TOTAL	SOURCE
2015	6	n/a	Innolitics / IntuitionLabs (Source: intuitionlabs.ai)
2022	91	n/a	IntuitionLabs (Source: intuitionlabs.ai)
2023	221	~690 (end 2023)	MDPI-sourced figure via IntuitionLabs (Source: intuitionlabs.ai)
Aug. 7, 2024 (interim)	107 (partial year)	950	MedTech Dive (Source: medtechdive.com)
2024 (full year)	253	n/a	IntuitionLabs / Innolitics (Source: intuitionlabs.ai)
2025	295 (record)	1,451 (end 2025)	The Imaging Wire (Source: theimagingwire.com)

Table 1 shows two distinct measurement snapshots (MedTech Dive's August 2024 interim count of 950 versus the full-year 253 figure reported later by Innolitics and IntuitionLabs), a discrepancy that reflects different collection dates rather than a data error, and this report presents both rather than silently reconciling them. The trend beneath the noise is unambiguous: annual clearances roughly doubled between 2022 and 2023 and have kept climbing since, while the underlying evidentiary bar for each individual clearance has not obviously risen in step. A JAMA Health Forum cross-sectional study of 691 FDA-cleared AI/ML devices through July 2023 found that decision summaries frequently omitted basic study information: study design was unreported for **323 devices (46.7%)**, training sample size for **368 (53.3%)**, and demographic composition for **95.5%** of devices, while only **6 devices (1.6%)** cited randomized clinical trial data and just **3 (under 1%)** reported actual patient health outcomes (Source: pmc.ncbi.nlm.nih.gov) (Source: pmc.ncbi.nlm.nih.gov). Postmarket surveillance data from the same study found premarket safety assessments documented for only **195 devices (28.2%)**, with **489** total adverse events reported across just **36 devices (5.2%)**, including one death, and **40 devices (5.8%)** recalled a combined **113** times, mostly for software defects (Source: pmc.ncbi.nlm.nih.gov) (Source: pmc.ncbi.nlm.nih.gov) (Source: pmc.ncbi.nlm.nih.gov).

UpDoc's own file sits comfortably within that pattern rather than outside it: no clinical study, human factors data in place of outcomes data, and safety controlled through deterministic-logic guardrails rather than trial evidence (Source: accessdata.fda.gov). IntuitionLabs' analysis of the FDA tracker adds a lifecycle dimension: as of the 2025 authorization cohort, "10% of AI/ML device clearances included PCCPs for iterative updates," signaling growing but still minority adoption of the pre-planned-change framework UpDoc used (Source: intuitionlabs.ai). Innolitics' 2025 year-in-review corroborates the figure precisely, noting that "a quarter of all devices were cleared in under 90 days, demonstrating the potential for a rapid path to market for well-prepared submissions" even as the median clearance time for the full cohort was 142 days (Source: innolitics.com).

Commercially, UpDoc's clearance arrives as investor and market-forecasting interest in clinical generative AI accelerates sharply. Grand View Research estimates the global generative AI in healthcare market at **$2.9 billion** in 2025, projecting growth to **$3.8 billion** in 2026 and **$28.2 billion** by 2033, a compound annual growth rate of **33.3%**, with North America holding a **41.0%** regional revenue share in 2025 (Source: grandviewresearch.com) (Source: grandviewresearch.com). UpDoc's **$18 million** seed round, while modest against that market backdrop, drew strategic capital from a drug manufacturer (Eli Lilly), a major health system (Mayo Clinic), and a patient-advocacy nonprofit investment arm (the American Diabetes Association's innovation fund), a combination that signals confidence in the regulatory strategy specifically, not just the underlying technology (Source: prnewswire.com). Broader hospital adoption of generative AI is already substantial: Grand View Research cites a JAMA Network study finding that "about 31.5% of U.S. hospitals had already implemented generative AI integrated with EHR systems by 2024, while nearly 25% planned implementation within the following year" (Source: grandviewresearch.com).

Case Studies and Real-World Examples

UpDoc, Inc.: K253281 and the December 2025 Clearance

UpDoc's own file remains the central case study for this report. FDA received the 510(k) on September 29, 2025 and issued clearance on December 23, 2025 for a device the agency's public database lists simply as "UpDoc V1.0," a prescription software medical device for insulin management in adults with type 2 diabetes (Source: innolitics.com). The company waited six months to disclose the clearance publicly, pairing the announcement with its **$18 million** seed round, an advisory role for former FDA Commissioner Dr. Robert Califf, and named pilot sites at **Cleveland Clinic**, **Allegheny Health Network**, and **UCSF Health** (Source: medicaldesignandoutsourcing.com) (Source: prnewswire.com). UCSF Health's chief pharmacy executive, Desi Kotis, framed the launch in continuity-of-care terms: "the next frontier is making care continuous with smarter coordination between visits," while Allegheny Health Network's primary care chair, Amy Crawford-Faucher, emphasized reduced "administrative burden" for physicians as the near-term value proposition (Source: prnewswire.com) (Source: prnewswire.com). Cleveland Clinic's Rohit Chandra reported no physician pushback and cited the FDA clearance itself as decisive to the pilot decision (Source: linkedin.com). Regulatory commentary on LinkedIn crystallized the central tension around the case within days: physician-engineer Ali Karimian wrote that "the FDA did not clear a Large Language Model. They cleared the cage built around it," a framing that Innolitics' formal analysis independently reached from the underlying FDA documents (Source: linkedin.com).

Digital Diagnostics' LumineticsCore: The Precedent for Autonomous AI Decisions

UpDoc is not the first FDA milestone in AI-enabled diagnosis or treatment support, and comparing it against the last genuine category-first clarifies how conservative UpDoc's clearance actually is. On April 12, 2018, FDA granted a **De Novo** request permitting the marketing of **LumineticsCore** (originally named IDx-DR), which the manufacturer describes as "the first autonomous, AI-based diagnostic system authorized for commercialization by the FDA," capable of diagnosing diabetic retinopathy "without the need for a clinician to also interpret the image or results" (Source: digitaldiagnostics.com) (Source: digitaldiagnostics.com). Company founder Michael Abramoff called the clearance "a historic moment that has the potential to launch a transformation in the way U.S. healthcare is delivered" (Source: digitaldiagnostics.com), while FDA's own contemporaneous

statement, delivered by Malvina Eydelman, director of the Division of Ophthalmic, and Ear, Nose and Throat Devices, noted that "today's decision permits the marketing of a novel artificial intelligence technology that can be used in a primary care doctor's office," and the product received expedited review through FDA's **Breakthrough Devices** program (Source: digitaldiagnostics.com) (Source: digitaldiagnostics.com). The contrast with UpDoc is precise: LumineticsCore genuinely made an autonomous clinical determination without physician review and required the harder De Novo pathway because no suitable predicate existed; UpDoc kept a physician in the dosing-decision loop and therefore could ride an existing predicate through the faster, cheaper 510(k) process.

Hygieia's d-Nav System: The Predicate That Made 510(k) Possible

UpDoc's predicate, Hygieia's **d-Nav System**, is itself worth treating as a discrete case study because its 2019 clearance quietly established the product-code and classification path UpDoc later reused. FDA's public 510(k) database lists d-Nav's clearance (K181916) as decided February 4, 2019, under device classification name "Calculator, Drug Dose," regulation number 868.1890, and indicates that no Predetermined Change Control Plan was authorized for the device, since the PCCP mechanism did not yet exist at the time (Source: accessdata.fda.gov). Innolitics' review of the predicate summary confirms that "d-Nav provides the next insulin-dose recommendation for adults with type 2 diabetes," using healthcare-provider-prescribed insulin instructions and glucose data, without any conversational or generative interface (Source: innolitics.com). Hygieia itself did not survive to see its predicate reused this way: the company filed for bankruptcy in 2024, roughly a year before UpDoc's own clearance was decided (Source: medicaldesignandoutsourcing.com). Six years and one generative AI wave later, that unremarkable dose-calculator clearance became the load-bearing precedent for what UpDoc, and its marketing team, would call a milestone in agentic clinical AI.

FDA's Digital Health Advisory Committee and the Unresolved Mental Health Chatbot Question

(Hypothetical Example within FDA's Deliberation): The fourth case is not a cleared device but an active regulatory deliberation that illustrates the outer boundary of what UpDoc's precedent does not yet resolve. At its November 2025 meeting, FDA's Digital Health Advisory Committee considered a hypothetical, not-yet-existing "therapy device" built on an LLM, evaluated across prescription and over-the-counter contexts, adult and adolescent populations, and single- versus multi-condition indications, explicitly to help the agency "clarify how regulation applies to medical devices based on newer forms of AI, including large language models that produce conversation-like outputs that are not predictable and may misguide users or lead to patient harm" (Source: statnews.com). The stakes of that deliberation are large: the FDA Law Blog noted that "currently there are 57.8 million adults who have a diagnosed mental illness and many do not have access to high quality and effective treatments," framing why generative AI mental health tools are commercially attractive even as their regulatory path remains unsettled (Source: thefdalawblog.com). Committee members reportedly considered guardrails including mandatory human oversight, transparency that outputs come from AI, restriction to prescription-only status, and even Class III designation for higher-risk configurations, while expressing particular concern about pediatric use, with the FDA Law Blog reporting that "for pediatric use, committee members were so concerned that they did not know what to say" (Source: thefdalawblog.com). No generative AI mental health device had been cleared as of that December 2025 summary (Source: thefdalawblog.com). Read alongside UpDoc, the DHAC meeting shows that FDA's comfort with patient-facing LLMs is highly use-case specific: bounded, provider-governed, physiologically deterministic tasks like insulin titration cleared via 510(k) within roughly twelve weeks, while open-ended, psychologically consequential conversational therapy remains an unresolved policy question even a full year later.

Implications and Future Directions

The UpDoc precedent is likely to accelerate a specific, narrow style of generative AI medical device submission rather than open the floodgates to autonomous AI clinicians. Innolitics author Yujan Shrestha's own prediction, published alongside his technical analysis, anticipates that "we will see more and more non-generative indications have conversational AI added on and get cleared through a 510(k) route," precisely the strategy UpDoc executed (Source: linkedin.com). As a regulatory consultant who has built a practice around exactly this pattern, Shrestha was more direct in his public reaction, describing the clearance as proof that "the market wants FDA-regulated LLM-enabled devices" and that "now there is one data point for how much it costs to get an LLM FDA cleared. Now we have a precedent" (Source: linkedin.com).

The bigger structural question is whether FDA's existing guidance stack, built primarily for locked or incrementally retrained ML classifiers, will hold up as more sponsors attempt the UpDoc pattern at scale. FDA's own January 2025 draft guidance solicited comment specifically on generative AI's fit within its lifecycle management framework, and as of this report's publication that guidance remains in draft form, with a comment period that closed April 7, 2025 and no finalized replacement yet issued (Source: fda.gov). Internationally, FDA has continued coordinating on machine learning lifecycle

principles with Health Canada and the UK's Medicines and Healthcare products Regulatory Agency, and in January 2025 the International Medical Device Regulators Forum released a 10-principle Good Machine Learning Practice document building on the original 2021 tri-agency principles, signaling that generative AI-specific international alignment, not just U.S. domestic guidance, is still being assembled (Source: fda.gov).

For life-sciences consultancies, health system informatics teams, and pharmaceutical manufacturers watching this space, several practical implications follow. First, the commercial appetite is real: Grand View Research's **33.3%** projected CAGR for generative AI in healthcare through 2033 implies that dozens of companies will attempt LLM-enabled SaMD submissions over the next several years, most without UpDoc's Stanford clinical-trial pedigree or its former FDA Commissioner advisor (Source: grandviewresearch.com) (Source: medicaldesignandoutsourcing.com). Second, health systems will increasingly need internal frameworks, separate from FDA clearance itself, for validating AI-supported interventions, defining escalation paths, and clarifying malpractice and liability exposure when an LLM-mediated workflow contributes to a clinical outcome, exactly the governance gap McGuireWoods flags for provider organizations deploying these tools (Source: mcguirewoods.com). Third, given how much of UpDoc's clearance depended on FDA's still-unfinalized generative AI lifecycle guidance and an evidence base that, per the JAMA Health Forum findings, remains thin industry-wide, near-term regulatory strategy for life sciences and AI teams should treat the 510(k)-predicate-plus-PCCP model as a proven but narrow template, not a universal solution, and should expect FDA's evidentiary expectations for less bounded, higher-autonomy generative AI use cases, mental health chatbots foremost among them, to tighten rather than loosen as the DHAC's deliberations mature into formal guidance (Source: pmc.ncbi.nlm.nih.gov).

Frequently Asked Questions (FAQs)

Is UpDoc really the first FDA-cleared LLM medical device? UpDoc is the first device FDA's public record confirms as a cleared SaMD using patient-facing large language models for data collection and communication, per the company's own framing, though the underlying dosing logic is deterministic, not generated by the LLM, and FDA's decision summary never uses the term "large language model" itself (Source: prnewswire.com) (Source: accessdata.fda.gov).

What is the UpDoc K253281 FDA clearance, exactly? It is a 510(k) substantial equivalence clearance, received September 29, 2025 and decided December 23, 2025, for prescription-only insulin-management software indicated for adults 18 and older with type 2 diabetes, using Hygieia's d-Nav System as its predicate (Source: innolitics.com) (Source: accessdata.fda.gov).

What is a 510(k) predicate strategy for an AI medical device? It means demonstrating that a new AI-enabled device shares the same intended use as an already-cleared, typically non-AI, predicate device, and that any technological differences do not raise new safety or effectiveness questions, allowing the sponsor to avoid a De Novo or PMA filing (Source: fda.gov).

What is a Predetermined Change Control Plan (PCCP)? A PCCP is a section of a marketing submission, finalized in FDA guidance on December 3, 2024, that pre-specifies future device modifications, their validation methodology, and their expected safety impact, so a manufacturer can implement authorized changes without submitting a new 510(k), De Novo, or PMA for each one (Source: mcdermottplus.com).

How does FDA regulate large language models in healthcare more broadly? FDA does not have a bespoke LLM or generative-AI classification; it applies existing SaMD frameworks (510(k), De Novo, PMA) case by case, supplemented by AI-specific guidance on PCCPs, Good Machine Learning Practice, and a still-draft lifecycle management guidance from January 2025 that explicitly seeks comment on generative AI's fit within that framework (Source: fda.gov) (Source: fda.gov).

How long does FDA clearance take for an AI medical device? Across the 2025 cohort of 295 AI/ML clearances, Innolitics reports that "the median time to clearance was 142 days, with an average of 150 days," and that "a quarter of all devices were cleared in under 90 days," while some more complex submissions took well over 200 days (Source: innolitics.com) (Source: innolitics.com).

Can a clinical AI system practice medicine autonomously after this clearance? No. Legal analysis of the clearance is consistent that "human-in-the-loop" and state-law practice-of-medicine requirements keep a licensed clinician accountable for treatment decisions, and UpDoc's own PCCP guardrails explicitly bar modifications from altering "core clinical decision-making" (Source: mcguirewoods.com) (Source: accessdata.fda.gov).

Conclusion

The plain regulatory record behind "**fda clears first llm as a medical device**" is narrower and more procedurally conventional than the June 2026 press cycle suggested. FDA cleared a prescription drug-dose calculator with a conversational data-capture layer, built on substantial equivalence to a six-year-old, non-conversational predicate, backed by human factors and software verification rather than clinical trial evidence, and constrained by a

Predetermined Change Control Plan that explicitly preserves deterministic dosing logic against future modification. That is a meaningfully different claim than "FDA authorized an autonomous AI clinician," and the gap between the two framings is exactly what generated the sharpest public debate around the clearance.

What the case demonstrates is durable and replicable: a well-defined predicate, a bounded intended use, disciplined separation between conversational interface and clinical decision logic, and early investment in a PCCP can move a generative AI healthcare product through the existing 510(k) system in a matter of months rather than years. For developers, health systems, and the life-sciences advisory firms supporting them, that template is now a validated starting point for AI SaMD strategy, not a guarantee that broader, more autonomous generative AI applications, particularly in higher-stakes domains such as mental health, will clear the same way. FDA's own Digital Health Advisory Committee was still actively debating the risk profile of generative AI mental health devices in the same season UpDoc's clearance became public, and no such device had been authorized as of this report's most recent verified sources. The UpDoc precedent opens a specific, well-lit path through FDA's regulatory system; it does not yet illuminate the far larger and more consequential question of how the agency will treat generative AI systems asked to reason, rather than merely converse, on a patient's behalf.

Tags: fda clears first llm as a medical device, updoc fda clearance, k253281, predetermined change control plan, 510k predicate strategy, generative ai samd, fda regulation of ai, clinical ai regulatory pathway, drug dose calculator predicate, de novo vs 510k

IntuitionLabs - Industry Leadership & Services

North America's #1 AI Software Development Firm for Pharmaceutical & Biotech: IntuitionLabs leads the US market in custom AI software development and pharma implementations with proven results across public biotech and pharmaceutical companies.

Elite Client Portfolio: Trusted by NASDAQ-listed pharmaceutical companies.

Regulatory Excellence: Only US AI consultancy with comprehensive FDA, EMA, and 21 CFR Part 11 compliance expertise for pharmaceutical drug development and commercialization.

Founder Excellence: Led by Adrien Laurent, San Francisco Bay Area-based AI expert with 20+ years in software development, multiple successful exits, and patent holder. Recognized as one of the top AI experts in the USA.

Custom AI Software Development: Build tailored pharmaceutical AI applications, custom CRMs, chatbots, and ERP systems with advanced analytics and regulatory compliance capabilities.

Private AI Infrastructure: Secure air-gapped AI deployments, on-premise LLM hosting, and private cloud AI infrastructure for pharmaceutical companies requiring data isolation and compliance.

Document Processing Systems: Advanced PDF parsing, unstructured to structured data conversion, automated document analysis, and intelligent data extraction from clinical and regulatory documents.

Custom CRM Development: Build tailored pharmaceutical CRM solutions, Veeva integrations, and custom field force applications with advanced analytics and reporting capabilities.

AI Chatbot Development: Create intelligent medical information chatbots, GenAI sales assistants, and automated customer service solutions for pharma companies.

Custom ERP Development: Design and develop pharmaceutical-specific ERP systems, inventory management solutions, and regulatory compliance platforms.

Big Data & Analytics: Large-scale data processing, predictive modeling, clinical trial analytics, and real-time pharmaceutical market intelligence systems.

Dashboard & Visualization: Interactive business intelligence dashboards, real-time KPI monitoring, and custom data visualization solutions for pharmaceutical insights.

Leading Claude & ChatGPT AI Enablement and Training: Help biotech and pharmaceutical teams adopt Claude and ChatGPT through practical training, governed workflows, use-case development, and implementation designed for regulated environments.

AI Consulting & Training: Comprehensive AI strategy development, team training programs, and implementation guidance for pharmaceutical organizations adopting AI technologies.

Contact founder Adrien Laurent and team at intuitionlabs.ai/contact for a consultation.

DISCLAIMER

The information contained in this document is provided for educational and informational purposes only. We make no representations or warranties of any kind, express or implied, about the completeness, accuracy, reliability, suitability, or availability of the information contained herein.

Any reliance you place on such information is strictly at your own risk. In no event will IntuitionLabs.ai or its representatives be liable for any loss or damage including without limitation, indirect or consequential loss or damage, or any loss or damage whatsoever arising from the use of information presented in this document.

This document may contain content generated with the assistance of artificial intelligence technologies. AI-generated content may contain errors, omissions, or inaccuracies. Readers are advised to independently verify any critical information before acting upon it.

All product names, logos, brands, trademarks, and registered trademarks mentioned in this document are the property of their respective owners. All company, product, and service names used in this document are for identification purposes only. Use of these names, logos, trademarks, and brands does not imply endorsement by the respective trademark holders.

IntuitionLabs.ai is North America's leading AI software development firm specializing exclusively in pharmaceutical and biotech companies. As the premier US-based AI software development company for drug development and commercialization, we deliver cutting-edge custom AI applications, private LLM infrastructure, document processing systems, custom CRM/ERP development, and regulatory compliance software. Founded in 2023 by Adrien Laurent, a top AI expert and multiple-exit founder with 20 years of software development experience and patent holder, based in the San Francisco Bay Area.

This document does not constitute professional or legal advice. For specific guidance related to your business needs, please consult with appropriate qualified professionals.