WHITE PAPER

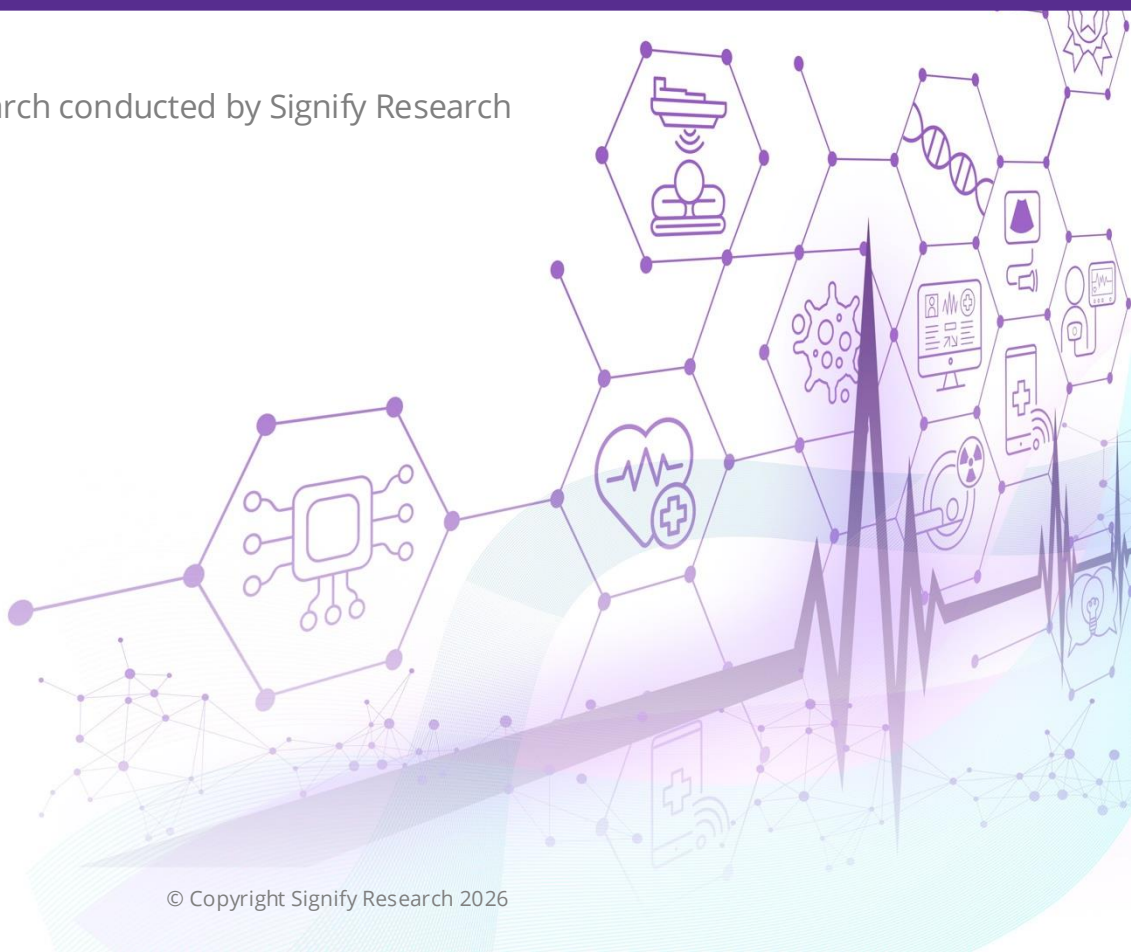
Signify RESEARCH

Voice AI in Specialty Patient Access

How Specialty Leaders Should Evaluate Solutions, Avoid Pitfalls, and Select the Right Partner

Based on primary research conducted by Signify Research

Published June 2026



0 Executive Summary

The specialty healthcare market is at an inflection point with Voice AI. Interest is high, vendor claims are abundant, and pilots are launching across the US. Yet a striking number of those pilots are stalling, not because the technology doesn't work, but because the criteria used to select them were poorly matched to operational reality.

This report draws on in-depth structured interviews with several US specialty group leaders spanning dermatology, ophthalmology, orthopedics, and multispecialty organizations, representing practices from two locations to 175 sites and 60 to 650 providers. Across every conversation, three tensions surfaced consistently:

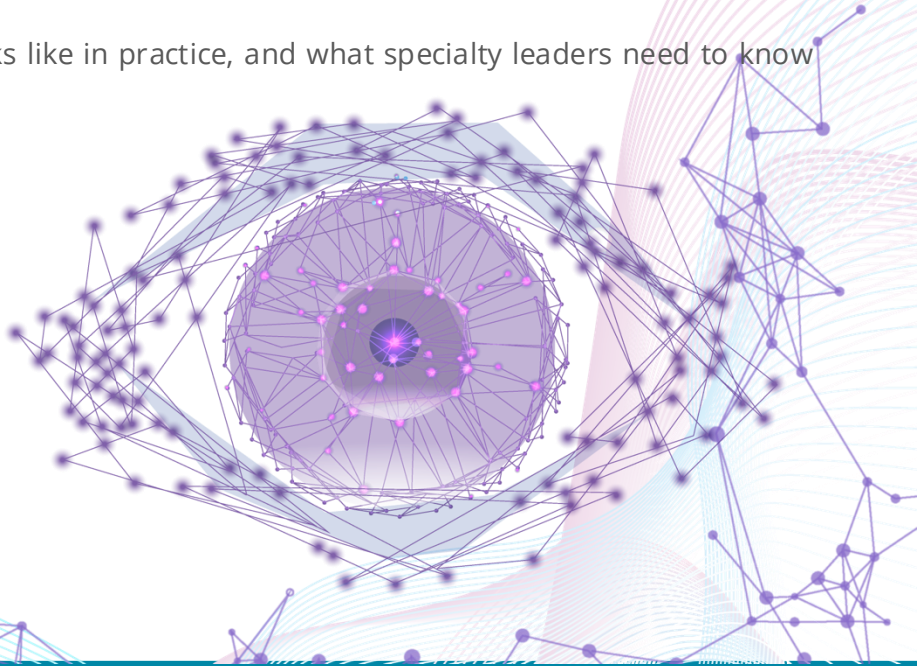
- They need to do more with less, and existing staffing models are no longer fit for purpose.
- They believe Voice AI is part of the answer, but most have no reliable way to validate vendor claims before committing.
- They are operating in a market where most solutions look identical on the surface, and the differences only become visible after go-live.

The interview findings point to a consistent conclusion. The organizations that succeed with Voice AI are not those that move fastest. They are those that evaluate most rigorously, define clear use cases, demand real-world validation, and select partners who can demonstrate sustained performance across an increasingly exacting set of operational and financial proof points.

The research points clearly toward what separates the deployments that hold from those that stall. The organizations making progress are selecting vendors against a consistent set of criteria: genuine EHR integration depth, not API access, but real-time write-back that reflects provider-specific rules under live conditions. Clinical safety logic built for healthcare environments, not retrofitted from other industries. Configuration flexibility that accommodates specialty workflows without requiring vendor involvement for every change. Reporting that unifies human and AI performance in a single view, so ROI can be demonstrated rather than asserted. And a vendor with the implementation rigour and financial stability to sustain performance beyond the pilot.

Vendors who can deliver across all of these dimensions in production, not just in demos, are rare. Finding them requires a more disciplined evaluation approach than most organizations currently apply.

This paper is built around what that looks like in practice, and what specialty leaders need to know before making a decision.



1 Why This Research, Why Now?

Specialty practices are not evaluating Voice AI because it is innovative. They are evaluating it because their operating model is under enduring structural pressure. Across every interview conducted for this research, four forces surfaced consistently, regardless of specialty or organizational scale.

Sustained high inbound call volume

Call volume has, in many cases, increased, while staff capacity to manage it has not kept pace. One senior director of operations at a 175-site multispecialty organization described the situation plainly: even with a centralized call center model, including offshore resources, the organization is still dealing with a fundamental mismatch between inbound demand and answering capacity. Calls overflow. Patients wait. Appointments are missed.

A particularly revealing finding from the research: at one multispecialty orthopedic organization with 26 locations, only 25 to 30 percent of inbound calls are for appointment scheduling. The remainder are requests for clinic hours, directions, medical records, billing inquiries, and general questions, all of which consume staff time without contributing to schedule utilization. This dynamic was echoed across multiple interviews: staff are spending significant capacity on non-scheduling calls, despite the primary operational objective remaining to fill appointments.

Staffing volatility and training burden

Hiring, onboarding, and retaining front-office staff remains one of the most operationally-disruptive challenges in specialty access. One dermatology group with 12 locations and 150 physicians reported front-desk and clerical staffing shortages and high turnover as primary drivers of the evaluation. Staff leave before they are fully trained. Coverage gaps create inconsistency in service delivery. Patients experience the variation.

One interviewee summarized the appeal of Voice AI in terms that resonated across the research: unlike staff, an AI system does not need breaks, vacations, or days off.

Rising patient expectations

Patients now expect immediacy. They compare their experience of calling a specialty practice to the responsiveness of consumer services, and the gap is increasingly visible in satisfaction metrics and online reviews. Several interviewees mentioned NPS and negative call center reviews as active performance concerns.

A crowded vendor landscape with unclear differentiation

The number of vendors claiming Voice AI capability in healthcare scheduling has grown rapidly. Most demonstrate well in controlled environments. Fewer perform consistently in live operations. One operations leader put it directly: the market is full of vendors who sound very similar, the difficulty is knowing which ones hold up when 200 calls come in on a Monday morning.

Beyond the operational pressure, patient access has evolved from an administrative function to a strategic one. It is now directly tied to revenue capture, schedule utilization, and patient retention, which means the stakes of getting Voice AI evaluation wrong have risen significantly.

2 What "Voice AI for Patient Access" Actually Means

One of the most consistent findings from the interviews is that specialty leaders do not yet share a common definition of what Voice AI for patient access means, or what it should be expected to do.

When asked directly, responses ranged from "answering phones and scheduling appointments" to a more sophisticated framing: a natural-language, human-sounding agent that can converse naturally with patients, answer questions, and interact with internal systems such as the EHR and CRM in the same way a human agent would.

One respondent, a senior operational leader at an ophthalmology platform with 110 locations, offered the clearest articulation: the expectation is that Voice AI should be a like-for-like swap for what human agents currently do. A full operational equivalent rather than a partial replacement or routing tool.

Most solutions currently on the market do not meet that bar, and the gap between the expectation and what vendors actually deliver in live operations is where most implementation disappointment originates. Understanding the gap between that aspiration and the current reality of Voice AI capabilities is essential for any organization beginning an evaluation.

In practice, Voice AI operates across two distinct layers of the patient access model, and the distinction has significant implications for how organizations should sequence adoption.

SEGMENT 1

Inbound Capabilities

- Answer and routing calls
- Scheduling, rescheduling , cancellations
- Triage and escalation logic
- Message capture and handoff to clinical staff

SEGMENT 2

Outbound Capabilities

- Appointment confirmations and reminders
- Recall campaigns
- No-show reduction
- Follow-up workflows

Two distinct layers of the access workflow

Inbound capabilities include answering and routing calls, scheduling and rescheduling, cancellations, triage and escalation logic, and message capture and handoff to clinical staff. These are operationally complex use cases. They require real-time interpretation of patient intent, tight integration with scheduling systems, and reliable handling of edge cases, including clinical urgency scenarios.

Outbound capabilities include appointment confirmations and reminders, recall campaigns, no-show reduction workflows, and structured follow-up sequences. These are comparatively more predictable, better understood by patients, and are lower risk from an implementation standpoint.

This distinction shapes how organizations should think about sequencing. Several interviewees described outbound automation as a lower-risk entry point, with faster time to visible value and lower patient resistance. One operations director noted that patients have been receiving automated reminders for years and largely expect them; the same psychological comfort does not yet exist for inbound AI.

> *"Inbound volume is constant in healthcare. Patients will always call. That's why inbound is more valuable, but also why we have to get it right."* - *Senior Director of Operations, 175-site multispecialty group*

3 The Four Challenge Clusters Specialty Practices Are Navigating

Across specialties and organization size, the operational challenges that motivate Voice AI consideration cluster reliably into four categories. Understanding these clusters is essential for defining evaluation criteria, because the solution requirements flow directly from the operational problems being solved.

Volume & Capacity Constraints

- Very high inbound call volumes, yet only 25–30% of inbound calls are actual scheduling requests
- Limited staff capacity
- Inconsistent answering at office level
- Call overflow into centralized call centers

Workflow & Scheduling Complexity

- Medication-dependent scheduling (injectable expiry windows, equipment constraints)
- Specialist workflow segmentation (cosmetic vs. medical vs. surgical)
- Appointment-type variability (new vs established vs procedure)
- Provider-imposed schedule restrictions
- Insurance verification burden
- Multi-location coordination

Patient Readiness & Variability

- Older patients less comfortable with automated tools
- Patients with specific accessibility needs, visual impairment, hearing difficulty, require system-level accommodation
- Insurance literacy challenges
- Language requirements in certain markets

Reliability & Operational Friction

- Integration errors create friction and erode staff trust in the system
- Latency or system inconsistency
- Messages not properly logged or routed
- Limited reporting visibility
- Escalation errors

Volume and Capacity Constraints

The most immediately visible driver of Voice AI interest is call volume. Practices with centralized call centers, multiple locations, or high per-provider appointment demand are managing a structural imbalance between patient demand and staff capacity. Call overflow, abandoned calls, and inconsistent answering at office level are symptoms of the same underlying problem.

In the research sample, this challenge appeared most acute in larger multispecialty organizations. One with 175 offices across nine states relies on a combination of onshore, offshore, and AI-enabled resources to manage inbound volume, and still describes conversion and access time as active improvement opportunities.

Workflow and Scheduling Complexity

Volume alone does not explain why Voice AI is difficult to deploy in specialty environments. The deeper challenge is scheduling complexity: provider-specific preferences, appointment-type variability, insurance verification burden, and multi-location coordination create workflows that are genuinely difficult to standardize or automate.

One orthopedic organization interviewed described a core scheduling challenge completely unrelated with call volume: routing patients to the correct specialty team, managing high procedural complexity, and avoiding misinformation from staff who were not specialized in the relevant workflow. Complexity, in their case, was a greater constraint than capacity.

In dermatology, the challenge takes a different form: short appointment slots of roughly ten minutes create high scheduling sensitivity, and the coexistence of medical, cosmetic, and surgical workflows, each with different payment pathways and provider preferences, requires routing logic that a generic Voice AI cannot replicate without specialty-specific configuration.

Patient Readiness and Variability

Across nearly every interview, patient demographics emerged as a meaningful constraint on automation. Older patient populations are less comfortable with automated voice interactions. Insurance literacy challenges create friction when patients try to navigate automated systems without understanding their own coverage. Language requirements in specific markets add additional configuration complexity.

One ophthalmology interviewee was direct: the retirement-heavy population they serve is a major barrier to inbound AI adoption. Patients who have vision issues and are calling about surgical procedures need a different kind of interaction than a younger, digitally-fluent patient scheduling a routine appointment. A Voice AI that cannot adapt to that variability will generate complaints.

Counterintuitively, one interviewee suggested that a well-designed AI might actually outperform some live agents in multilingual or clarity-sensitive contexts, noting that patients sometimes struggle to understand certain live agents, and that an AI with clear, slow speech and multilingual support could work better for some demographics.

Reliability and Operational Friction

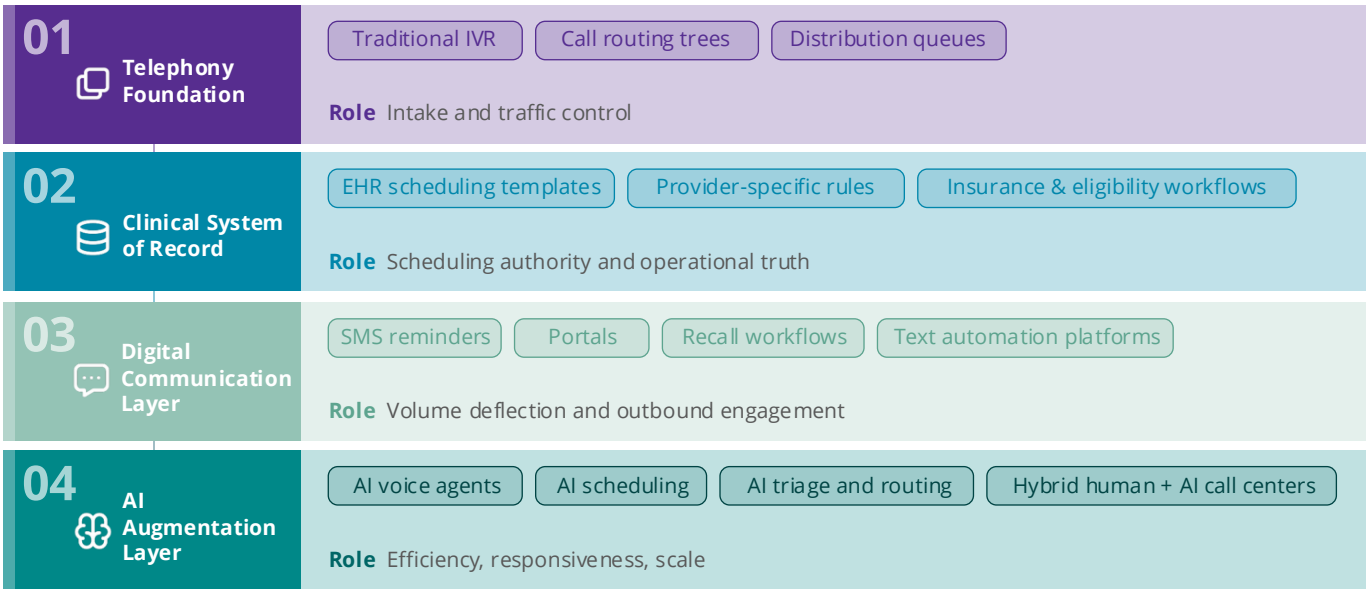
A fourth cluster of challenges relates to the performance of existing systems, including early Voice AI deployments. Latency, system inconsistency, messages not properly logged or routed, limited reporting visibility, and escalation errors are not theoretical risks. Several interviewees described these as active pain points with their current technology stack.

One respondent when describing their legacy solution described latency as "a little long," infrastructure as inconsistent, and scalability as lacking, and is actively evaluating alternatives. Another flagged that the primary limitation of their current system is the inability to customize routing logic to reflect their specific operational model.

> *"The AI worked in the demo. The real question is whether it works when 200 calls come in on a Monday morning."* - *Operations leader, multispecialty organization*

4 Adoption Reality: Layering, Not Replacing

A critical misconception in how Voice AI is often discussed is the implication that it replaces existing infrastructure. It does not, at least not in the near-term. Specialty practices are integrating Voice AI into a layered access stack that already includes telephony systems, EHR scheduling platforms, digital communication tools, and in some cases, centralized call centers. Understanding this architecture is essential for understanding where implementation risk lives.



The research confirms that specialty practices are not replacing this stack. They are integrating AI into it, and the success of that integration depends entirely on how well the AI layer connects to the other layers. Several interviewees described using multiple platforms simultaneously because no single system covers all workflows: one for outbound reminders, one for referral automation, one for scheduling, one for phone routing. This fragmentation reflects the genuine complexity of the environment rather than being a failure of technology strategy.

One respondent organization described a stack that includes Promptly for referral automation, Solutionreach for reminders, Brevium for recall, RingCentral for telephony, and Nextech as the EHR, plus an AI scribe for documentation. Their observation was direct: no single platform covers all workflows effectively. Each solves a different operational need.

This layered reality creates four structural pressures that every Voice AI deployment must resolve.

1. Does the AI reflect real-time scheduling templates and provider-specific rules, or is it working from stale or incomplete data?
2. Are complex, clinical, or atypical scenarios handed off safely and immediately, with no ambiguity in responsibility?
3. Are human and AI performance metrics unified into a single operational view, enabling attribution clarity?
4. And can measurable improvement be isolated and defended across channels?

Organizations that underestimate these pressures during vendor selection tend to encounter them at the worst possible moment, during go-live.

5 Where Most Evaluations Go Wrong

Across interviews, a consistent set of gaps emerged in how specialty groups evaluate Voice AI solutions. These gaps rarely surface during vendor selection, but reliably surface during implementation, often at significant operational and financial cost.

The demo problem

Vendor demonstrations are optimized for clean scenarios and predictable inputs. Live patient access environments introduce noise, emotional complexity, insurance ambiguity, and edge cases that demos are not designed to expose. Several interviewees described the gap between demo performance and live performance as one of their primary concerns when evaluating new vendors.

One respondent currently using an AI-enabled phone system flagged a specific incident: a patient asked about combining medications, and the system provided incorrect guidance due to a lack of clinical understanding. This was a clinical safety incident enabled by a system that could not recognize the limits of its competence. This type of edge case is exactly what demos do not surface, and exactly what escalation logic is supposed to prevent.

> *"I'm sceptical that case studies match real-world outcomes. I want to know what happens when the system encounters something it wasn't trained for."*
>
> *- C-suite executive, multispecialty center*

Features versus workflow depth

Many solutions can demonstrate isolated capabilities convincingly. Fewer can handle the full complexity of a real scheduling scenario, including provider-specific rules, visit type variations, insurance eligibility, and exception handling, without requiring manual correction or escalation.

One of the clearest articulations of this gap came from an interviewee at a musculoskeletal organization whose scheduling model is built almost entirely around clinical procedures rather than standard consultations. Injectable treatments with pharmaceutical expiry windows, equipment and location constraints, provider credentialing requirements, and strict follow-up cadences for conditions like fractures and chronic pain management create a scheduling environment where the margin for error is narrow. A Voice AI deployed without configuration that reflects all of these constraints will not merely create administrative friction, it will book appointments that cannot be safely fulfilled.

Integration assumed, not verified

Integration is frequently discussed in terms of API availability. The operational reality of interacting with EHR scheduling systems, particularly for write actions like booking, is far more nuanced. Multiple interviewees flagged EHR integration as a top-three requirement, and several described the lack of real-time EHR sync as a non-starter.

One respondent currently using tier 1 ambulatory EHR vendor noted that their AI phone system is not integrated with the EHR at all, meaning every call handled by AI still requires a manual step to complete the scheduling action. The telephony layer and the scheduling layer are effectively disconnected. This is not an unusual situation in the current market; it is a common architectural gap that vendors frequently understate during evaluation.

Change management underestimated

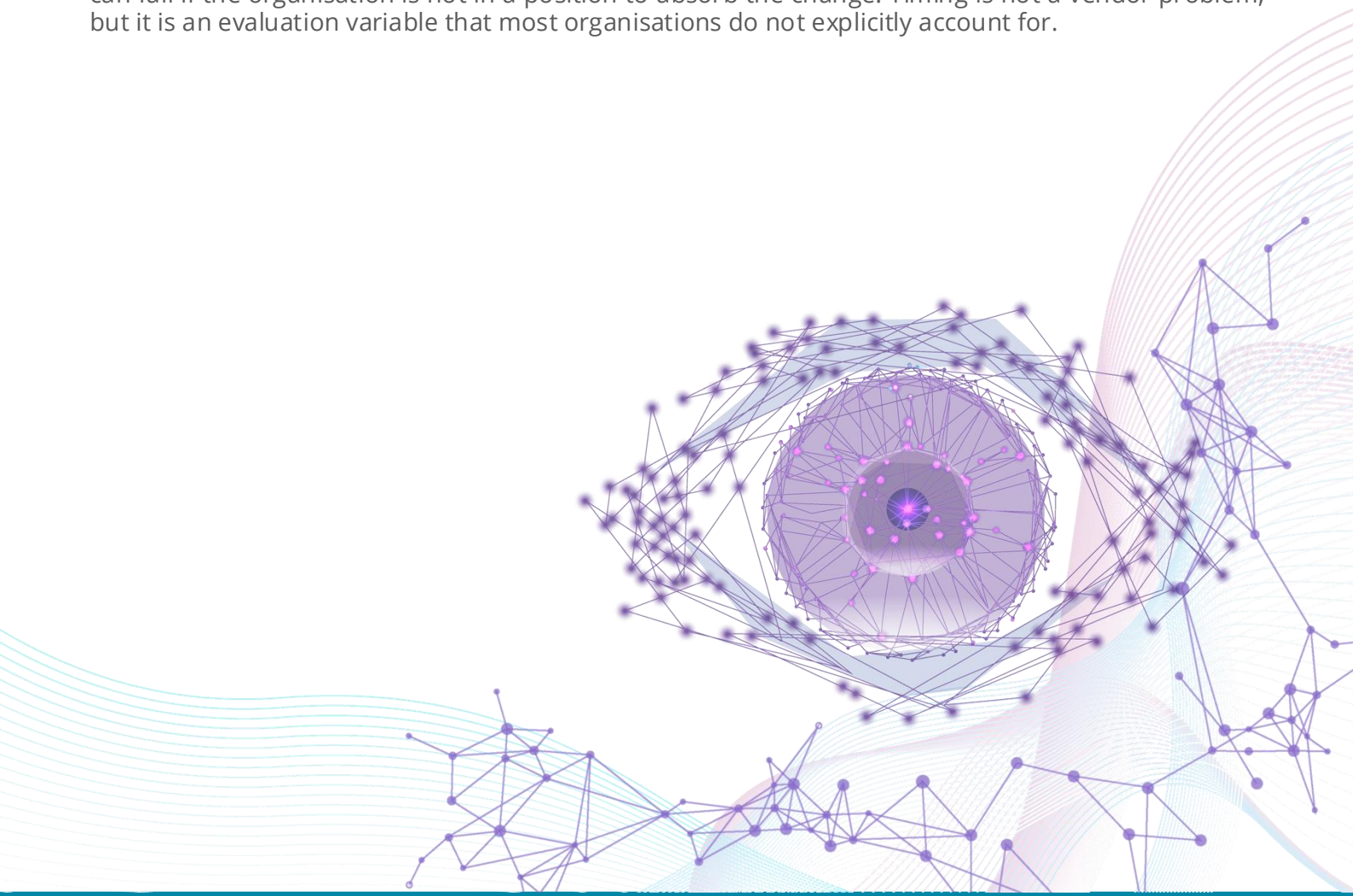
Even technically sound implementations can stall if staff do not trust the system, if workflows are not redesigned to accommodate it, or if the metrics used to evaluate success are not aligned with the new operating model. Several interviewees raised staff acceptance as a meaningful implementation risk, not because staff are resistant to technology, but because they are appropriately sceptical of systems that create more work through errors, routing failures, or escalation gaps.

One respondent was notably direct: staff will be upset if the implementation reduces headcount. Leadership needs to position Voice AI honestly, not as a supplement to staff when the actual intent is a reduction in FTEs. The mismatch between how AI is positioned internally and what it is expected to do is one of the more predictable sources of adoption failure.

Organizational timing overlooked

A gap that rarely appears in vendor evaluations but consistently affects outcomes is organisational readiness at the point of deployment. Voice AI evaluation does not happen in isolation, it competes for IT bandwidth, governance attention, and implementation capacity alongside EHR migrations, acquisition integrations, and other infrastructure projects.

Several interviewees explicitly deferred Voice AI decisions until competing priorities cleared, with one organisation stating that no new technology commitments would be made until their EHR migration was complete. The practical implication is that even a well-chosen vendor and a well-designed pilot can fail if the organisation is not in a position to absorb the change. Timing is not a vendor problem, but it is an evaluation variable that most organisations do not explicitly account for.



6 How Performance Should Actually Be Measured

One of the clearest signals of organizational maturity in Voice AI adoption is whether a practice has defined what success looks like before deployment begins, and whether it has the measurement infrastructure to validate it afterward.

Across the research, a wide range of sophistication in performance tracking was evident. The most advanced organization in the sample, a 175-site multispecialty group, tracks overall call volume, calls handled, calls handled per hour, conversion rate, average transaction time, repeat caller rates, call abandonment rates, unavailable time for reps, and NPS, and reviews these metrics weekly. Most smaller groups track far fewer metrics, and several acknowledged tracking almost nothing formally.

Performance measurement clusters into four dimensions, each connected to a different layer of operational and financial accountability.

Volume & Throughput Efficiency

How much work is moving through the system?

- Total call volume
- Calls handled per hour
- Average transaction time
- Repeat caller rate
- Unavailable time / rep utilization

Smaller groups often lack the reporting infrastructure to baseline their own performance, making ROI validation harder before and after deployment

Access & Conversion Outcomes

Are patient calls successfully turning into visits?

- Conversion rate (call → booked appointment)
- Seen-to-scheduled rate
- Appointment fill rate
- No-show and cancellation rates

New patient target: same-day or next-day access. Established patients: 3–5 days depending on speciality

Responsiveness & Service Level

How quickly and reliably are patients helped?

- Call abandonment rate
- Wait time
- Time to schedule
- Return-call timing

Most smaller speciality groups cannot tell you their call abandonment rate, even when they cite it as a top concern

Experience & Quality Signals

Is the interaction strengthening or weakening patient trust?

- NPS
- Complaint volume
- Negative call centre reviews
- Escalation accuracy

Patient experience metrics carry revenue implications, yet are commonly monitored only in larger enterprise settings.

A notable finding from the research is the measurement gap at smaller organizations. Several practices acknowledged that they do not formally track call abandonment rates or time-to-schedule, despite identifying responsiveness as a top-three priority. This gap is itself a risk factor: organizations that cannot baseline their current performance cannot validate improvement. Vendors who cannot support this level of attribution will find it increasingly difficult to sustain contracts beyond an initial pilot.

The direction of travel in Voice AI evaluation is clear. Organizations are moving away from asking whether AI reduced workload, toward demanding attributable, measurable lift across all four dimensions, defensible at the executive and board level.

> *"The main KPI is 100% resolution of all calls and tasks within 24 hours. We track missed calls, answer time, and escalation rates. Not just whether the AI handled the call, but whether the patient's problem was actually solved."* - *Director of Operations, musculoskeletal MSO*

7 The Deployment Proof Path

Voice AI investments do not deliver value all at once. They move through a progression of operational and financial proof points, and failure at any stage tends to halt adoption entirely. The research surfaced a consistent pattern in how organizations evaluate success over time.

Understanding this progression is critical for both buyers and vendors. Organizations that expect immediate, fully-realized ROI will be disappointed. Those that structure their deployment as a staged validation process, with clear success criteria at each stage, are significantly more likely to achieve long-term adoption.

The ROI Credibility Ladder

Early Signal (0–3 Months)

Question: Is anything moving?

- Reduction in abandonment
- Early conversion lift
- Reduced call pressure
- Positive staff feedback
- In outbound conversion use cases, early signals may be visible within weeks

Failure Mode: If no visible early lift → pilot stalls.

Operational Stability (3–6 Months)

Question: Is it reliable?

- Scheduling accuracy holding
- Escalation working
- Integration stable
- No major patient complaints

Failure Mode: Integration errors erode trust quickly.

Financial Defensibility (6–12 Months)

Question: Can we defend this at board level?

- Staffing impact quantified
- Cost savings vs TCO validated
- Performance lift attributable
- Clear ROI narrative

Failure Mode: Attribution ambiguity weakens CFO support

Strategic Sustainability (12+ Months)

Question: Is this structurally improving our model?

- Stable performance over time
- No degradation in experience
- Scalable across specialities/sites
- Governance comfort with reporting, including audit trails, escalation logs, and compliance documentation

Failure Mode: Performance volatility or reporting gaps trigger reassessment.

The research confirmed that expectations around time-to-value vary significantly by organization type. The more operationally-sophisticated organizations, those with dedicated analytics teams and weekly performance reviews, expect meaningful signals within six months and a defensible ROI narrative within twelve. Smaller organizations sometimes have shorter timelines: one respondent expected to see results within three months and defined success as reduced escalations, fewer complaints, and improved routing accuracy.

One interviewee at a surgical ophthalmology organization articulated the most demanding version of this expectation: for outbound conversion use cases, they would expect to see improvement in consult conversion within one to two weeks of launch. Their reasoning was direct: if the AI does not handle the call, staff still must handle it and no value is created.

8 What a High-Performing Voice AI Access Model Looks Like

Based on the research, a mature Voice AI-enabled patient access model delivers consistently across seven measurable dimensions. These are not aspirational capabilities, but are the requirements that the research identified as both necessary for sustained adoption and frequently absent in current deployments.

Scheduling Integrity at Scale

Real-time alignment with provider-specific rules, visit types, insurance constraints, and multi-site variability, without manual correction or downstream clean-up.

Escalation Safety by Design

Clear containment boundaries and frictionless handoff for clinical, urgent, or atypical scenarios, with no ambiguity in ownership or responsibility.

Defensible Performance Lift

Sustained, attributable improvement in conversion rates, abandonment reduction, staffing efficiency, and schedule optimisation, measurable under executive review.

Experience Stability Across Demographics

Consistent interaction quality across age groups, communication channels, and levels of patient sophistication, without erosion of trust or increased complaints.

Unified Operational Intelligence

Integrated reporting across AI and human workflows, enabling executive-level visibility, attribution clarity, and continuous optimisation.

Configuration Agility

The ability to evolve scheduling logic, provider templates, escalation pathways, and specialty workflows over time.

Adoption Durability

Maintaining sustained staff adoption and patient acceptance across all deployment phases.

Most Voice AI solutions can demonstrate strength in one or two of these dimensions during a controlled evaluation. Very few can demonstrate consistent performance across all seven in live operations.

9 A Practical Evaluation Framework

Based on the research, the following five-step framework reflects how the most rigorous organizations are approaching Voice AI vendor selection. It is designed to move evaluations beyond surface-level feature comparisons toward operational reality.

Step 1: Define the use case explicitly

Organizations should be specific about whether they are addressing inbound call volume, outbound engagement, a specific metric such as conversion or abandonment rate, or a combination. Without explicit use case definition, vendor evaluation becomes unfocused and comparisons become meaningless. The research showed that organizations with the clearest use case definitions tended to reach vendor decisions more efficiently and experience fewer post-implementation surprises.

Step 2: Assess capabilities against the seven dimensions

Rather than evaluating against a feature checklist, use the seven dimensions of a high-performing access model as the evaluation rubric. Ask vendors to demonstrate, not just describe, performance against each dimension. Escalation logic and scheduling accuracy under realistic edge cases should be tested explicitly.

Step 3: Validate through a controlled pilot

Leading organizations do not rely solely on demonstrations. They structure pilots around a narrow, well-defined use case, run them over a defined period, and measure against clearly established baseline metrics. The research sample included organizations that had completed pilots ranging from 30 days for a surgical outbound use case to six months for inbound scheduling. The pilot is not a proof of concept; it is a proof of operational fit.

Step 4: Apply the ROI Credibility Ladder

Define what success looks like at 90 days, six months, and twelve months before the pilot begins, and ensure the vendor is accountable to the same milestones. Organizations that define ROI criteria after deployment tend to find that vendor and buyer expectations diverge in ways that are difficult to resolve.

Step 5: Verify integration depth, not API availability

Specifically test write-back capability into your EHR scheduling system under realistic conditions, including edge cases, provider-specific rule enforcement, and multi-location scenarios. EHR integration is the most common point of failure in live deployments and the most important capability to validate before signing. Several interviewees cited lack of real-time EHR sync as a non-starter, and one described an active deployment where the telephony and scheduling layers remain disconnected, requiring manual steps to complete every AI-handled interaction.

10 Vendor Evaluation Checklist

The following checklist is derived directly from the research findings. It is designed to be used during vendor presentations, pilot design, and contract negotiation.

Criteria	Yes	Partial	No
Integration & Technical Requirements ★ Must-Pass			
Real-time read & write EHR integration, not just data access?	○	○	○
Integration demonstrated with your specific EHR platform and version?	○	○	○
Provider-specific scheduling rules and appointment-type logic supported?	○	○	○
Direct access to patient insurance and eligibility data to inform scheduling?	○	○	○
EHR latency profile acceptable for patient-facing interactions?	○	○	○
System behavior during EHR downtime clearly defined and safe?	○	○	○
Escalation & Clinical Safety ★ Must-Pass			
Escalation triggers from AI to human agent configurable?	○	○	○
Clinical urgency signals, such as symptoms mentioned, detected and routed?	○	○	○
Full audit trail for all escalations available?	○	○	○
AI fails gracefully when it does not know the answer?	○	○	○
Specialty Configuration			
Existing deployments in your specific specialty?	○	○	○
System configurable for your appointment types and provider preferences?	○	○	○
Cosmetic, medical, and surgical workflows differentiated and routed correctly?	○	○	○
Process defined for updating scheduling logic when provider preferences change?	○	○	○
Patient Experience & Demographics			
System performs well with older or less digitally confident patients?	○	○	○
Multilingual capabilities supported and appropriate for your patient population?	○	○	○
Emotionally charged or frustrated callers handled appropriately?	○	○	○
Human agent option always available and configurable in the call flow?	○	○	○
Performance Reporting & Attribution			
Granular metrics reported at rep, site, specialty, and channel level?	○	○	○
Human and AI performance available in a unified operational view?	○	○	○
Attribution modeling for conversion, abandonment, and access time?	○	○	○
Reporting data accessible quickly and to appropriate staff?	○	○	○
Vendor Assessment			
Documented healthcare-specific deployment methodology in place?	○	○	○
Implementation timeline and team requirements clearly defined?	○	○	○
Support SLA and escalation process documented?	○	○	○
References from same-specialty organisations at similar scale available?	○	○	○
Financial stability and product roadmap evidenced?	○	○	○

Conclusion

Voice AI is no longer an experimental technology in specialty patient access. It is becoming a foundational layer of how specialty practices manage demand, optimize operations, and engage patients. The organizations represented in this research are actively deploying, piloting, or planning deployments, and the pace of adoption is accelerating.

But the gap between potential and realized value remains wide. The market is crowded with vendors making similar claims. Most of them demonstrate well. Fewer deliver consistently in live specialty environments, where scheduling complexity, EHR integration depth, escalation logic, and demographic variability create conditions that controlled demonstrations are not designed to expose.

The organizations that succeed are those that approach adoption with operational discipline: defining clear use cases, demanding real-world validation, structuring deployment as a staged proof process, and holding vendors accountable to measurable outcomes at every stage.

The market will consolidate around solutions that can demonstrate performance across the full spectrum, from scheduling integrity and escalation safety to defensible ROI and experience stability across patient demographics. For specialty leaders, the goal should not be to adopt Voice AI quickly. It should be to adopt it in a way that actually holds up.

About this research

This report is based on primary research conducted by Signify Research, drawing on in-depth structured interviews with specialty group leaders across the United States. Interviewees represent organizations ranging from 2 to 175 sites and 60 to 650 providers, spanning dermatology, ophthalmology, multispecialty, and orthopedics. Buyer roles represented include operations, IT, and executive leadership, including C-suite decision-makers.

Interviews were structured around a consistent set of topics covering access KPIs, inbound versus outbound value, integration requirements, risk thresholds, and buying journey and ROI expectations. All quotes and characterizations are anonymized or composite representations of research findings.

About Artera

Artera is an agentic company strengthening how healthcare providers communicate and care for patients. As an agentic partner, we bring over a decade of healthcare experience to address urgent workflows from day one and build custom solutions as healthcare providers' needs evolve. Trusted by 1,000+ specialties, FQHCs, health systems, and federal agencies, Artera strengthens and enhances patient relationships across every interaction - from intake and scheduling to referral management, post-visit care, and more.

About Signify Research

Signify Research is a specialist HealthTech market intelligence firm, providing data-driven insights that help vendors, investors, and health system leaders make better strategic decisions. Headquartered in the UK and serving clients across North America, Europe, and Asia-Pacific, Signify Research combines primary research, including in-depth interviews with technology vendors and healthcare professionals, with sales data reported directly by leading vendors, to deliver the most complete and authoritative view of global HealthTech markets.

Coverage spans five core areas: Medical Imaging, Clinical Care, Digital Health, Diagnostics and Lifesciences, and Healthcare IT, with dedicated analyst expertise across AI in Healthcare. Clients benefit from direct access to specialist analysts, off-the-shelf market reports and subscriptions, and custom research services tailored to specific strategic needs.

Signify Research's methodology is independent and impartial, underpinned by a commitment to quality, transparency, and actionable intelligence.

About the author

Vlad joined Signify Research in 2023 as a Senior Market Analyst in the Digital Health team. He brings a strong background in strategy consulting and market intelligence, with experience leading commercial strategy, market planning, and due diligence projects for governments, operators, healthcare providers, and technology vendors.

At Signify Research, Vlad works closely with clients on custom research and advisory engagements across the digital health landscape, helping organisations navigate evolving market dynamics, competitive positioning, and growth opportunities. His work combines rigorous market analysis with a practical understanding of how healthcare technology is adopted and scaled in real-world settings.