ChatGPT

Strategic Roadmap: Agentic AI for Contextual Next Best Actions in Oncology

Introduction

Oncology brands face complex challenges with **prescription switching** – when healthcare providers (HCPs) change patients' treatments due to new data, competitive actions, or patient factors. Capturing and responding to these switching patterns in real-time is crucial for commercial success. Data-driven approaches have shown promise (e.g. one AI-driven switch prediction model identified new HCP targets and boosted *new-to-brand* prescriptions by 80% ①), but current Next Best Action (NBA) systems in pharma are often limited to static rules or black-box algorithms. This strategic roadmap outlines a phased implementation of an **Agentic AI** solution that combines large language models (LLMs), autonomous agents, rule-based guards, and reinforcement learning (RL) to deliver highly contextual NBAs in the U.S. oncology market. The roadmap is divided into four phases, each building capability and confidence: - **Phase 1: Enrichment Layer** – Add explainability and HCP-specific summaries using LLMs. - **Phase 2: Agent Copilot** – Introduce an agentic AI assistant to support Territory Account Manager (TAM) workflows through multi-step planning and orchestration. - **Phase 3: Root Cause Engine** – Integrate structured and unstructured data to diagnose why HCPs switch therapies, providing deep insights into switching patterns. - **Phase 4: Autonomous Optimization** – Evolve to a self-optimizing system that combines causal machine learning, RL, and agent self-evaluation to adapt NBA recommendations continuously.

Each phase is detailed with its goals, technical components, data requirements, risks with mitigations, and differentiators. Throughout, we emphasize where LLMs provide unique value (e.g. reasoning, narrative generation, decision justification) and where they must defer to rule-based or RL systems (e.g. compliance enforcement, precise optimization, scalability). This phased approach ensures **feasibility and trust**, starting with enhancements that deliver immediate value and progressing toward a robust Agentic AI platform for commercial optimization in oncology.

Phase 1: Enrichment Layer – Add Explainability and HCP Summaries via LLMs

Goals & Value Delivered: The goal in Phase 1 is to enrich existing NBA outputs with human-readable explanations and HCP context summaries. This "explainability layer" increases TAMs' trust in and adoption of suggestions by clarifying *why* an action is recommended and providing relevant background on each HCP. In sensitive domains like healthcare, stakeholders demand transparency and justification for AI-driven recommendations ② . By summarizing complex data into a narrative, the AI will help TAMs quickly grasp each HCP's situation (e.g. treatment history, recent switches, key patient factors) and the rationale behind the next-best action. This delivers immediate value: better-informed sales reps who can approach conversations with greater context and confidence.

Technical Components: Phase 1 introduces a **Large Language Model** into the workflow as a generative explanation and summarization engine. The LLM (e.g. GPT-4 or a domain-tuned model) is prompted with an HCP's relevant data – such as recent prescribing trends, prior interactions, patient mix – and the recommended action (from the existing NBA engine). It then produces two key outputs: (1) a concise **HCP profile summary** highlighting pertinent details (e.g. "Dr. Smith has increased use of Competitor X in renal cell carcinoma since Q2, possibly due to new clinical data on efficacy in younger patients.") and (2) an **explanation of the NBA suggestion** (e.g. "Next Best Action: Schedule a meeting – recommended because Dr. Smith's recent switch to Competitor X may indicate questions about our therapy's dosing; a meeting can address these concerns."). The LLM's chain-of-thought reasoning ability is leveraged to connect data points into a coherent reason [3], while ensuring it stays grounded in the actual data provided (to avoid hallucinations). This component is orchestrated such that the LLM's output is appended to the existing CRM/NBA interface (e.g. as a tooltip or expandable note next to each suggestion).

Data & Integration Requirements: Key data inputs include the **structured data** already in use by the NBA system (HCP profiles, sales and prescription data, prior suggestions) and relevant **unstructured data** such as call notes or HCP engagement history. In Phase 1, integration focuses on extracting these data points from CRM and data warehouses and feeding them into the LLM prompt. A *data ingestion and preprocessing pipeline* may vectorize text (e.g. rep call notes) and retrieve recent pertinent info for each HCP [4] [5]. The LLM can be accessed via a secure API or on-premise if privacy is a concern. Importantly, the system must incorporate **compliance filters** on the LLM's output – e.g. ensuring no off-label claims or policy-violating language is generated. This can be done by rule-based checks on the generated text and by carefully constraining the prompt (for instance, instructing the LLM to only use provided data and approved messaging).

Risks & Mitigations: The primary risks in this phase are *LLM accuracy* and *compliance*. LLMs are prone to **hallucination**, i.e. generating information not supported by data. In a pharma context this could mean a wrong assumption about why a switch happened, which would erode trust. Mitigations include using **retrieval-augmented generation** (providing the model with factual data snippets to ground its answer) and testing the outputs thoroughly. Notably, recent best practices address hallucinations by grounding LLM outputs in real reference data [6]. Another risk is **inappropriate or non-compliant content** – the LLM might produce explanations that, while sensible, use language against regulations (e.g. mentioning a use case beyond the product label). To mitigate this, Phase 1 uses a limited generation scope: the LLM mainly summarizes known data and reasons within approved messaging guidelines. Deterministic **rules** will guard the final output (for example, a rule to flag and remove any mention of unapproved indications). Additionally, all LLM outputs can be logged for review to ensure they meet compliance and quality standards before fully deploying to field users. *Latency* is a minor consideration – generating a summary per HCP will add processing time, but since it's an offline or asynchronous enrichment (e.g. generated overnight or a few minutes before a call), it's manageable. If needed, smaller distilled models or prompt optimizations can be applied to improve speed [7].

Differentiators: Compared to current NBA approaches, which might offer a raw data insight or a cryptic score as justification, this enrichment layer provides **clear, narrative justifications** for actions. It transforms the system from a "black box" into a **transparent advisor**, which is crucial for user adoption and regulatory comfort [2]. Competitors' systems often lack such explainability or rely on static rule-based messages. Here, by leveraging an LLM's natural language generation, we deliver **personalized, HCP-specific context** at scale – something that rules alone cannot do. This phase uses LLMs to do what they excel at: **reasoning across disparate data and generating a coherent explanation** [8]. However, unlike end-to-end generative systems, our approach limits the LLM to an advisory role; all critical recommendations still originate from existing vetted algorithms, ensuring precision. Phase 1 sets the foundation for more advanced agentic functionality by establishing trust and understanding – TAMs begin to see the AI as a partner that not only suggests actions but also *explains its thinking*, a key first step in the Agentic AI journey.

Phase 2: Agent Copilot – Introduce Agentic Planning to Support TAM Workflows

Goals & Value Delivered: Phase 2 aims to elevate the system from a passive suggestion generator to an **active AI assistant (Copilot)** for the sales rep (TAM). The Agent Copilot will support multi-step planning and task orchestration, acting as a digital partner that can handle complex queries and proactive planning. The core value delivered is **workflow augmentation**: helping TAMs prioritize and execute their daily activities in a more strategic, data-driven way. For example, rather than just showing independent suggestions ("Email Dr. Jones", "Schedule lunch with Dr. Lee"), the agent can synthesize a plan: "This morning, focus on Dr. Jones and Dr. Lee. Dr. Jones recently switched 3 patients away from our drug – I have drafted an email explaining recent supporting data. Afterwards, schedule a lunch with Dr. Lee next week to discuss her patient feedback." Such multi-step orchestration saves the rep time and ensures no opportunity is missed. The Agent Copilot can also answer ad-hoc questions (e.g. "Why did Dr. Smith reduce use of our drug?" or "Show me top 5 priority HCPs for this quarter and why") by dynamically analyzing data and leveraging the reasoning capabilities introduced in Phase 1. Ultimately, this phase delivers **greater efficiency and effectiveness** – TAMs can cover more ground with less guesswork, and the AI continuously adapts plans to real-world changes (like a true partner). It begins to realize the promise of Agentic AI: an autonomous problem-solver that perceives, decides, and acts in an evolving environment [9] .

Technical Components: The centerpiece of Phase 2 is an **Agentic AI Orchestration layer** built around an autonomous agent (or a set of agents) that interfaces with both the LLM and existing tools. Technically, this could be implemented using an agent framework (e.g. LangChain or Microsoft Autogen) that enables the AI to perform a perception-reasoning-action loop [10] [11] . Key components include: - **LLM-based Planner:** The agent uses an LLM to interpret high-level goals or questions from the rep and to break them into actionable steps. The LLM's reasoning (possibly guided by a ReAct prompt style) allows the agent to plan a sequence of actions rather than single-step responses [11] [12] . For instance, if the goal is to "improve Dr. Smith's engagement," the agent might decide: step 1 – analyze Dr. Smith's recent activities; step 2 – if low interaction, draft a personalized email; step 3 – set a reminder to follow up in 2 weeks and flag if no response. - **Tool Integrations:** The agent is equipped with connectors (APIs or function calls) to enterprise systems: CRM (to log activities or read interaction history), email scheduling tools, sales analytics dashboards, etc. It can **execute actions** like sending a suggestion to an email tool, booking a meeting on the TAM's calendar (with confirmation), or pulling up a detailed report, through API calls [13] . - **Rule-Based Governance:** A set of **business rules and compliance checks** is embedded in the agent's decision loop to serve as guardrails. For example, rules will prevent the agent from scheduling more than X meetings per day or contacting an HCP too frequently (to comply with engagement policies). The agent's logic incorporates these constraints when planning (this is akin to the *symbolic AI hybrid approach* where rule-based logic is combined with probabilistic reasoning for decisions [14]). - **Context and Memory:** The agent maintains state – context about the TAM's territory, ongoing engagements, and past outcomes – to avoid redundant or conflicting actions. A vector database or internal memory may store summaries of past interactions so the agent remembers context beyond a single turn [15] . For example, if Dr. Lee prefers email over calls, the agent remembers this preference in future suggestions. - **Multi-Agent or Hierarchical Setup (if needed):** As complexity grows, the system can be structured into specialized sub-agents. For instance, a *Planning Agent* handles global scheduling and prioritization, while an *Insight Agent* handles on-demand analysis questions, and a *Compliance Agent* oversees that any proposed action passes rules (similar to having agents specializing in different aspects like planning, execution, and error recovery [14]). These agents can communicate to ensure the final plan is optimal and safe. This modular approach increases reliability, as each agent can be simpler and focused.

Data & Integration Requirements: In Phase 2, **real-time data integration** becomes more important. The Agent Copilot will need access to up-to-date information on HCP activities (sales data, recent switches, engagement history) and possibly external data like formulary changes or medical news (if within scope). Data pipelines may be set to refresh key metrics daily or in near-real-time so the agent's decisions reflect the latest state. Integrations to **calendar and email systems** are needed if the agent is to schedule meetings or draft communications. The LLM planner will use the data from Phase 1's enriched context (summaries, explanations) as one input, along with any new data signals introduced in Phase 2. For example, if an HCP's formulary status changed (structured data) or if there's a new publication affecting the product (unstructured data), those should feed in. Technically, an event-driven architecture could trigger the agent – e.g. a new data point like "Dr. Smith just had 2 new switches to Competitor X" could prompt the agent to recompute suggestions for Dr. Smith. **APIs** (or function calls in the LLM) will connect the agent to the CRM to log recommendations or retrieve account details on the fly. Importantly, all integrations must maintain **HIPAA and data privacy compliance** – patient-level data should be aggregated or de-identified when used, and HCP identifiers kept secure. Phase 2 may require working closely with IT to establish secure access for the agent to perform actions on behalf of users, with audit logs.

Risks & Mitigations: With the agent taking on more autonomous functions, **reliability and control** are key concerns. One risk is the agent making suboptimal or **erratic planning decisions** – e.g. suggesting too many actions, or focusing on the wrong priorities – especially in early stages or if the LLM misinterprets the goal. To mitigate this, the agent should initially act in a *recommendation/assist mode* rather than full autonomy. The TAM remains in control: the agent might draft an email or propose a schedule, but the human approves or adjusts it. This "human-in-the-loop" approach catches mistakes and also helps train the agent (through feedback on which suggestions were accepted). Another risk is **compliance/privacy breaches** if the agent mishandles data or suggests actions beyond allowed marketing practices. This is mitigated by the robust rule constraints – essentially a built-in **"Constitutional AI" layer that ensures compliance with business rules** [3] . For example, if an agent's plan somehow tries to suggest an off-label discussion, a rule would intercept and modify or block that step. Performance and **latency** could also be a challenge: the agent's planning loop involves multiple LLM calls and tool calls. We address this by optimizing the agent's reasoning scope (keeping prompts focused and using efficient models for straightforward tasks). Caching of results and limiting how far ahead the agent plans can also help keep response times reasonable for the user. Lastly, **user experience** risk: TAMs might be overwhelmed if the agent over-communicates. Mitigation: design the Copilot interface to be intuitive, allowing the rep to ask for help when needed or to view the agent's plan in a digestible format (perhaps a daily brief). Gradual onboarding and training for the field team will be essential so they learn to work effectively with the AI assistant.

Differentiators: The Agent Copilot sets our solution apart from traditional CRM suggestions by providing an **interactive, intelligent assistant** rather than static recommendations. Current approaches typically rely on a list of next-best-actions ranked by a model – they don't help the rep decide *how* to execute those actions or *how to coordinate multiple actions* across their territory. In contrast, our agent brings **autonomy and proactivity**: it can handle complex queries and adjust plans on the fly, something rules-based systems cannot do. For example, a conventional system might flag HCPs A, B, C as priorities. The agent goes further to say: “Focus on A and B today for these reasons, and defer C to next week when new data will be available” – a level of contextual planning only possible with multi-step reasoning. This phase also differentiates by **augmenting the TAM’s capabilities**: it’s like giving each rep a personal analyst and coordinator who never forgets any detail. Competitors beginning to experiment with AI assistants might still lack the integration of *both* advanced planning and strict compliance checking. Our use of an **ensemble (LLM + rules + tools)** ensures the agent’s recommendations are not only smart but also actionable and within guardrails – a critical advantage in pharma. In sum, Phase 2 moves the solution from decision support toward **collaborative automation**, where the AI works hand-in-hand with the human rep. It leverages LLM strengths for flexible reasoning and language (e.g. answering “why” questions or drafting messages) while leaning on deterministic systems for precision execution and compliance (e.g. automatically logging activities, enforcing contact frequency limits). This balance is a key competitive differentiator, marrying creativity with control in the sales workflow.

Phase 3: Root Cause Engine – Integrate Structured & Unstructured Data to Find Why HCPs Switch

Goals & Value Delivered: Phase 3 focuses on understanding the *why* behind prescription switching patterns through a **Root Cause Engine**. The goal is to unearth deep insights into HCP behavior – specifically, to determine the drivers of switching (e.g. switching a patient from our therapy to a competitor’s, or vice versa). By integrating diverse data sources and applying advanced AI analysis, this engine will answer questions like: *Are HCPs switching due to efficacy concerns, side effect profiles, patient demographics, formulary changes, or competitor promotions?* The value of knowing the reasons behind switches is enormous: it allows tailoring of next-best actions not just to *that* a switch occurred, but to the **cause** of the switch. For instance, if the engine finds that Dr. Allen is switching patients due to **toxicity concerns**, the next-best action might be to share new safety data or patient management tips; whereas if Dr. Allen’s switches are driven by an **aggressive competitor rebate**, the action might be to involve an access support team. By finding root causes, the organization can move from reactive suggestions to **truly contextual, targeted interventions**. This phase delivers value to multiple stakeholders: TAMs get richer insights to guide their customer conversations (“I see you’ve had issues with patient insurance approvals – let’s discuss our support program”), sales managers can refine strategy (focusing on the highest-impact switch drivers), and marketing teams can craft resources to address common switch reasons. In essence, Phase 3 strengthens the *insightfulness* of the AI – making it not just a recommender, but also an analyst that explains HCP behavior.

Technical Components: The Root Cause Engine is a composite of data analytics, machine learning, and LLM-driven text analysis: - **Data Fusion Layer:** First, a wide array of data sources must be integrated. **Structured data** includes prescription trends, patient claims (de-identified, aggregated by HCP), formulary & insurance data, sales calls and email engagement metrics, and any measurable trigger events (e.g. a competitor drug approval date, guideline changes). **Unstructured data** includes free-text sources: field notes entered by reps, transcripts of advisory boards or physician surveys, medical literature or news feeds, and possibly HCP social media or forum posts (if accessible and compliant). This phase likely involves setting up a data lake or warehouse that links these sources by HCP or geography and time. The system might use a knowledge graph or simply an index that aligns events with behavior changes. - **Analytical**

Models: On the structured side, traditional **ML models** or statistical analysis can look for correlations and potential causal links. For example, a time-series analysis or changepoint detection model could flag that "Dr. Allen's decrease in Drug A prescriptions correlates with the approval of Competitor B in April". More advanced, a causal inference model (using techniques like difference-in-differences or causal forests) might be employed to test hypotheses about what factor is likely causing the switch. On the unstructured side, **NLP models** (potentially powered by LLMs) will extract themes and sentiments: e.g. scanning rep notes for mentions of "side effects" or "efficacy concerns", or reading a medical news feed to see if Dr. Allen's region had a lot of exposure to a new study. LLMs can be used to summarize these unstructured corpora and even generate hypotheses. For instance, the LLM might be prompted with: *"Given Dr. Allen's prescription patterns and these notes/articles, what are possible reasons for switching from Drug A to Drug B?"* Modern LLMs can produce plausible hypotheses ("Possible causes: The physician may have seen new clinical trial results favoring Drug B's efficacy in a subgroup, or had patients experiencing side effects with Drug A..."). These **LLM-generated hypotheses** can then be cross-checked against data (this is similar to using LLMs for insight generation [16] [17]). - **Causal Reasoning Agent:** We can extend the Phase 2 agent to a specialized **Root Cause Agent** that orchestrates the above analyses. This agent would take an HCP or cohort as input and perform a series of steps: query the data for any major changes (e.g. drop in our drug's share), then retrieve potential explanatory data (like competitor activities, notes), then call an LLM to synthesize a narrative explanation. The output is a **Root Cause Analysis report** for that HCP (or segment of HCPs) which highlights likely causes for switches, supporting evidence, and confidence level. This can be delivered to TAMs as an on-demand explanation ("Why might Dr. Allen be switching?") or proactively when significant switches are detected. - **User Feedback Loop:** Incorporating user knowledge is valuable – reps often have anecdotal insights (e.g. "Dr. Allen told me many of her patients struggled with our injection form"). The engine can allow TAMs to input feedback, which the agent can treat as additional data to refine its analysis for that HCP. Over time, this could even be used to **fine-tune the LLM or adjust model weights** if certain root causes repeatedly prove true or false.

Data & Integration Requirements: This phase requires broad **data integration** efforts. Access to patient-level data (even if aggregated) is sensitive – we must ensure compliance with HIPAA and other privacy regulations. Likely, we will use de-identified patient-switch data from third-party sources (e.g. claims vendors) where we can see how many patients of Dr. X switched therapies and when. Integration with medical information sources is also needed: e.g. connecting to a medical literature database or competitive intelligence feed to know what external events could drive switching. A **text analytics pipeline** will ingest rep notes, medical news, etc., possibly using an LLM with **NLP capabilities to label or summarize texts** (identifying if a note mentions a side effect, or if a news article is about a new indication). We also need an **analytic environment** where data scientists can implement and update the causal inference models – likely using Python/R with libraries for causal ML. This environment should be linked to the Root Cause Agent so that the agent can trigger running these analyses on demand. All this must integrate back into the CRM or analytics dashboard for consumption: e.g. a field in the CRM that shows "Likely Switch Reason: Efficacy concern (confidence 80%)" with an explanation. From an infrastructure view, Phase 3 may involve building a **unified data model** keyed by HCP and time, merging data from different silos. This is a non-trivial but foundational requirement for both this phase and the next.

Risks & Mitigations: The pursuit of causal insights comes with risks around **accuracy and interpretability**. One risk is that the engine might identify spurious correlations as root causes (e.g. linking a switch to an irrelevant event). To mitigate this, we combine algorithmic rigor with human validation. The causal ML models will be validated on historical cases (does it correctly find known causes of past switches?) and we will limit the engine to proposing causes that have at least some plausible support. The LLM might sometimes **hallucinate a cause** that sounds reasonable but isn't backed by data. Mitigation: use **retrieval augmentation and grounding** – any narrative the LLM provides should reference specific data points or documents (for example, cite the specific study or note it is referring to), ensuring traceability. Another risk is **data quality and completeness**: if the data sources are incomplete, the engine could miss key causes. For instance, if we lack data on a competitor's patient support program, the engine won't flag cost issues even if that's the real driver. Addressing this requires continually expanding data sources and acknowledging uncertainty. We might have the engine express a confidence level and even say "data insufficient" if it can't find clear drivers. There's also a **user comprehension risk** – the output might be complex. We mitigate this by making the root cause report *narrative and visual*: e.g. a short paragraph explanation (LLM-generated) accompanied by a simple chart or table of supporting data, so the TAM can digest it quickly. Lastly, **compliance** is a concern: internally discussing why a doctor prefers a competitor can be sensitive. We must ensure these insights are used appropriately (for internal strategy, not as direct messaging to the HCP unless handled tactfully). The agent and TAM should use root cause insights to inform their approach, but always in a compliant manner (e.g. if cost is an issue, the rep knows to talk about access programs rather than disparaging the competitor). Proper training and guidelines around using these insights will mitigate misuse.

Differentiators: The Root Cause Engine is a significant leap beyond what most current pharma NBA systems offer. Traditional systems might flag that an HCP's behavior changed, but they rarely explain **why**. Our solution will uniquely provide a *because* – illuminating the drivers behind HCP decisions. This is a differentiator on multiple levels: - **Depth of Insight:** Instead of just identifying targets or actions, we deliver understanding. For example, a competitor's platform might use ML to predict who is likely to switch, but we are providing the rationale (e.g. "likely to switch **because** of X"), giving our sales and marketing teams a chance to preempt or counteract it. This is akin to moving from predictive analytics to **prescriptive/ diagnostic analytics**. - **Combining Unstructured with Structured Data:** Many systems stick to quantitative metrics. By leveraging LLMs to interpret unstructured text (emails, notes, publications), we tap into a rich vein of information that others might ignore [8]. This means our recommendations account for things like sentiment and qualitative feedback, which is hard to quantify but often crucial in oncology decisions. - **Causal, Not Just Correlative:** We emphasize understanding causality – a step beyond typical AI which often finds correlations. By incorporating causal inference techniques and expert domain knowledge (possibly through the LLM or knowledge graphs), our engine strives to distinguish true causation from coincidence. This reduces false leads and focuses reps on what truly matters. - **Closed-loop Learning:** The insights from this engine will also feed back into the system. For example, if a certain cause is identified frequently (say many HCPs switching due to a competitor's new data), the organization can respond broadly (adjust strategy or provide new educational materials). Over time, this learning loop can even inform the RL in Phase 4 about which types of actions succeed against certain root causes. - In summary, Phase 3 differentiates our approach by adding **intelligence and narrative** to the raw analytics. It transforms a flood of data into a coherent story about each HCP. This not only helps TAMs tactically but also gives leadership a strategic view of market dynamics. It's a capability that competitors will find hard to replicate quickly because it requires both breadth of data and advanced AI integration. We are essentially building a domain-specific "investigative AI" that becomes smarter with each switch observed, positioning our organization ahead in understanding the oncology market's shifts.

Phase 4: Autonomous Optimization – Combine Causal ML, RL, and Agent Self-Evaluation for Adaptive NBA

Goals & Value Delivered: Phase 4 represents the fruition of the Agentic AI vision – an **autonomously optimizing NBA system** that continuously learns and improves itself. The goal is to have a solution that not only reacts to the environment but *proactively evolves* its strategy to maximize outcomes (e.g. increased adoption of our therapy, better HCP engagement, improved patient outcomes) within ethical and compliance boundaries. Concretely, this means the system will use live feedback to refine its recommendations and even try novel tactics in a controlled way, functioning as an always-on "brain" for commercial optimization. The value delivered is an unprecedented level of adaptability and efficiency: as the market changes (new competitors, shifting treatment paradigms, seasonality), the AI adapts in near-real time, without waiting for the next model retraining cycle. It will also personalize at a granular level – each HCP can have a dynamically optimized engagement plan, moving beyond one-size-fits-all segmentation. By combining **causal learning** (understanding what truly works and why) with **reinforcement learning** (experimenting and learning from results), the system aims to recommend the *best possible action* at any given moment, not just based on historical patterns but on an evolving policy that improves with experience. Ultimately, Phase 4 delivers a **self-optimizing, empathetic commercial engine** that can handle complexity and scale beyond human limitations [18]. It reduces the need for constant manual tuning of strategies and allows the business to respond quickly to new opportunities or challenges – a significant competitive edge in the fast-moving oncology landscape.

Technical Components: Phase 4 integrates several advanced components: - **Reinforcement Learning (RL) Framework:** The backbone of autonomous optimization is an RL system that treats the commercial engagement process as a sequential decision-making problem. The state can be defined by the HCP's profile and current context (prescribing status, engagement history, external environment), and actions are the possible next best actions (call, email with specific content, invite to event, etc.). A **reward function** must be crafted – for example, a reward could be tied to positive outcomes like increase in prescribing (on a lag), HCP engagement (meeting scheduled or response), or even an intermediate proxy like a survey of HCP intent. The RL agent (or agents) will use algorithms (e.g. deep Q-networks, policy gradients, or bandit optimization) to learn which sequences of actions yield the best long-term results. Importantly, given the relatively slow and sparse rewards in pharma (prescription changes take time), we might use *simulation* or *batch RL* on historical data initially. The RL agent can be first trained on retrospective data (learning what actions historically led to better outcomes) and then continue to learn online. - **Causal ML Integration:** One innovation is to use the **causal insights** from Phase 3 to inform the RL policy. Rather than learning from scratch, the system can bias exploration toward actions that address identified causes. For example, if the Root Cause Engine indicates an HCP is switching due to efficacy concerns, the RL agent's policy can weight educational actions higher for that HCP. In effect, the causal model can act as a prior or context for the RL. This can accelerate learning and make the RL more sample-efficient by reducing random trial-and-error. Additionally, causal techniques can be used to continually evaluate *policy changes* – ensuring that we attribute outcomes correctly to the AI's actions (did sales go up because of the agent's recommendation or due to unrelated market forces?). This helps in refining the reward function over time. - **Agent Self-Evaluation (Meta-Reasoning):** Beyond the pure numeric optimization of RL, the Agentic AI system can include an LLM-driven **self-evaluation mechanism**. After each cycle of recommendations (say after a week or a campaign), the agent can analyze its own performance using the LLM to identify patterns or errors. For instance, it might generate a natural language report: "Over the last month, actions involving educational content had the highest success in converting switches, whereas increased call frequency did not yield improvement [19] [20]. The agent should prioritize quality of engagement over quantity." This kind of reflection can be done by prompting the LLM with logs of actions and outcomes, letting it summarize and even suggest strategy tweaks. The agent can then adjust its strategy parameters or suggest a new experiment, effectively creating a **closed feedback loop** where qualitative insight complements quantitative learning [21] . - **Multi-Agent Competition and Collaboration:** At this mature stage, we may deploy multiple agents with different objectives that **compete or collaborate** to find optimal solutions [3] . For example, one agent could focus on maximizing short-term engagement, another on long-term relationship quality; or more concretely, an agent representing product A and another for product B (if we have multiple brands) might negotiate the best approach for a shared HCP. A *governance agent* might monitor these interactions to ensure compliance and fairness (no one agent "oversteps" promotional boundaries). This echoes concepts from multi-agent systems where specialized agents optimize different reward signals and the ensemble achieves a balanced outcome [3] . - **Automation & Execution Layer:** In Phase 4, the AI system could be allowed to trigger certain actions autonomously (with oversight). For instance, it might auto-send an approved email when an opportunity is detected or adjust the cadence of outreach without asking every time. The execution layer from Phase 2 is extended to allow **algorithmic tweaks in near-real-time** – e.g. automatically reallocating reps' time towards HCPs where the agent predicts higher return, or personalizing content selection using generative models. All such autonomous actions would still be logged and available for human review, and critical actions (like changing call plans drastically) could still require approval depending on comfort level. The key is that the system *operates continuously*, monitoring data, updating recommendations, and sometimes acting on its own within predefined limits.

Data & Integration Requirements: Phase 4 requires closing the loop on data: a robust **feedback capture mechanism**. We need to gather outcome data for each recommended action: Did the TAM execute it? If yes, what happened (meeting occurred, any change in Rx)? If no, why not (TAM overruled it, HCP refused, etc.)? This implies tighter CRM integration to mark which suggestions were taken and what the results were. Potentially, integrate with sales outcome data (weekly new prescriptions, etc.) to attribute outcomes to actions. Real-time or frequent updates are needed so the RL agent gets timely reward signals. Also, if running online experiments (A/B testing different strategies), the system must track those and feed results back. On the computational side, the RL training/inference might need scalable infrastructure (possibly cloud-based with GPU acceleration if using deep learning). The LLM usage in self-evaluation could be heavy, but that might be done offline (e.g. generate a weekly strategy report). Integration with compliance monitoring systems is also key – as the system becomes more autonomous, any violations or near-misses should trigger alerts. Essentially, Phase 4 ties together all prior data integrations and adds the requirement of **continuous monitoring** of both performance and compliance. We might employ a dashboard for AI performance, showing KPIs like acceptance rate of suggestions, success rate, etc., which draws from the integrated data.

Risks & Mitigations: An autonomously learning system in a regulated environment comes with serious risks that must be managed: - **Reward Misoptimization:** The RL agent might find a way to "game" the reward that isn't truly aligned with business goals or ethics (the classic alignment problem). For instance, if the reward is primarily based on Rx count, the agent might inadvertently learn a strategy that over-focuses on high-volume prescribers at the expense of new patient identification or, worse, it could suggest aggressive tactics that risk relationships or compliance. Mitigation: design a **robust reward function** that incorporates multiple facets (engagement quality, compliance adherence, patient outcomes) and involve domain experts in reviewing it. Also put hard constraints in the agent (rules that cannot be violated even if the reward might seem to encourage it). Frequent audits of the agent's suggestions by a human panel can catch any drift in undesirable directions. - **Overfitting and Instability:** The agent might overfit to short-term trends or noise, causing oscillating strategies (“thrashing”). Mitigation: enforce a degree of exploration and require statistical significance before the agent fully commits to a strategy change. We can also use smoothing techniques and maintain a baseline policy to fall back on if the new policy underperforms. Essentially, use A/B testing or champion-challenger setups where the new RL-driven policy is tested against the last known good policy before full rollout. - **Safety and Compliance:** The more autonomous the system, the more we need failsafes. We will maintain the **rule-based “safety layer”** that overrides any action violating compliance or business rules, no matter what the RL suggests [3]. For example, if the RL agent for some reason suggests an off-label promotion (maybe it saw a reward in some data because a rep went off script and it worked), the compliance rules must stop that outright. Additionally, we likely will *not* allow the agent to execute irreversible actions without human confirmation – e.g. it can recommend to send an email, maybe even draft it, but sending to an HCP might still need a quick human review in the UI (or at least use only pre-approved content). - **Transparency and Trust:** As the system becomes complex, explaining why it’s suggesting what it does becomes harder (this is a known challenge with advanced AI). We mitigate this by continuing to use the LLM to provide **natural language explanations** for its decisions at runtime, via chain-of-thought or stored rationale [22]. Also, governance tools can monitor for bias or unusual patterns (ensuring the AI isn’t inadvertently favoring or excluding certain HCP groups unfairly). By 2025, regulations might even demand audit trails for AI decisions; our design from Phase 1 onward, with logged explanations, positions us well to provide that transparency. - **Technical Complexity:** Combining RL, causal models, and LLMs is cutting-edge and complex. There is risk in the feasibility and maintenance of such a system. To mitigate, we plan a gradual progression (hence the phases) and will possibly pilot the fully autonomous features in a controlled region or product line first. This way, we can learn and adjust with limited exposure. Moreover, we keep a **human governance committee** in place throughout Phase 4 to oversee the AI’s learning and ensure it aligns with strategic goals and ethical practice.

Differentiators: If successfully implemented, Phase 4 yields a solution that few (if any) competitors will have: a **self-learning, self-improving commercial AI agent**. This is the ultimate differentiator – it’s like having a million experiments running and learning from each other to find the best engagement approach, whereas others rely on annual model updates or static playbooks. Key differentiators include: - **Continuous Adaptation:** Traditional systems require manual re-calibration when market conditions change (e.g. a new competitor launch might require new rules or re-running models). Our agent will detect shifts (say, a drop in performance of certain actions) and adapt its policy autonomously, in near-real time. This could be the difference between retaining HCPs during a competitor launch or losing them – an adaptive system reacts faster than human-driven cycles. - **Personalization at Scale:** Through RL, the system can learn subtle preferences of each HCP (effectively modeling a policy per HCP or micro-segment). Where a human marketer might create 5 segments, the AI might find 50 nuanced clusters or truly individualized approaches. And it does so while still respecting constraints and common sense, thanks to the integrated reasoning. This one-to-one optimization is a big competitive advantage in a market where relationships matter. - **Synergy of Techniques:** Our ensemble approach – Agentic AI, LLM, rules, RL, causal analysis – means we cover all bases. LLMs bring **contextual understanding and narrative** abilities that pure numeric models lack, such as explaining an action in plain language or creatively brainstorming new tactics [3]. RL contributes **evidence-based optimization**: it won’t just guess an action; it will have tried variants and seen the results. Rules ensure **consistency and compliance** throughout. Competitors might have one or two of these, but not all. For example, a competitor might implement an RL for NBA (some advanced pharma companies do bandit-based suggestion engines), but without LLMs they lack explainability and flexibility; or they might use a fancy LLM chatbot for reps that sounds smart but isn’t learning from outcomes systematically. We combine the strengths of both. As an industry-first, we would have an agent that can explain itself (“why this action was chosen”) and prove itself (“and I have data to back that it works”) – a powerful combination. - **Empathetic and Ethical AI:** By Phase 4, our system isn't just maximizing a metric blindly. Thanks to the design, it factors in quality and ethical constraints (for instance, it could deliberately avoid pushing too hard on a physician who is overwhelmed – an aspect of empathy – because it recognizes long-term trust is more valuable). The inclusion of an explicit governance layer and possibly even encoded ethical principles means our AI works in a way that aligns with human values and pharma regulations, which is a differentiator in gaining trust from internal and external stakeholders. In a sense, our AI becomes a responsible agent on the team, not just a number-cruncher. - In summary, Phase 4 delivers an AI-driven commercial engine that is **dynamic, intelligent, and self-correcting**. It transforms NBA from a static recommendation list into a living strategy that co-evolves with the market. This is the capstone of our roadmap – turning the vision of an Agentic AI optimizing oncology engagement into reality, and setting a new standard for the industry.

Trade-offs, Feasibility, and Limitations

Implementing this Agentic AI roadmap requires careful consideration of trade-offs at each step. We have deliberately structured it in phases to manage **feasibility and risk**, but it is important to remain self-critical about the current limitations of GenAI and agentic systems:

- **LLMs: Unique Strengths vs. Limitations.** Large language models bring powerful capabilities in understanding and generating natural language, making sense of unstructured data, and simulating reasoning. They shine in tasks like summarizing HCP profiles, generating explanations, and providing a conversational interface for the agent [8]. These are areas where deterministic systems fall short, and leveraging LLMs creates a more human-like, engaging AI. However, LLMs are not sources of ground truth – they can produce incorrect or fabricated information if not properly directed. They also lack inherent knowledge of specific numeric business rules or constraints. Therefore, in our design LLMs are **augmented and checked by other components**. For example, the LLM might draft a personalized message, but a rule system ensures it's compliant; or the LLM might propose a reason for a switch, but the causal model verifies it. We accept that current LLMs (as of 2025) may occasionally err or require heavy compute, so we use them where their value exceeds their risk: in making the system smarter and more user-friendly, but *not* as the final arbiter of critical decisions or calculations. This interplay ensures that we capture LLMs' unique value (contextual reasoning and language generation) while **deferring to rules or RL for precision and reliability** [3].

- **Rules and Deterministic Logic: Scope and Scalability.** We rely on rule-based components for compliance and business logic, because some constraints are non-negotiable (e.g. legal regulations, or policies like "don't contact an HCP more than 3 times in a week"). Rules are excellent for enforcing hard boundaries and providing consistency. They are also transparent – easy to explain to regulators and easy to adjust when policies change. However, a pure rule-based system is brittle and doesn't scale well in complexity (the number of rules can explode, and interactions become hard to manage). Our approach uses rules as a **safety net and framework**, not as the driver of content. The trade-off is that while rules will constrain the AI (possibly preventing some creative strategies), they greatly reduce risk. Over time, we might encode some rules into the AI's "constitution" so that even the LLM or RL agent intrinsically avoids those actions [3]. This hybrid ensures we don't end up in a scenario where the AI suggests something that a human would immediately know is inappropriate.

- **Reinforcement Learning and Causal ML: Data Challenges.** RL is powerful for optimization, but its effectiveness depends on having sufficient experimentation and clear feedback. In pharma sales, cycles are long and data is noisy; an RL agent might struggle to get enough signal quickly. We mitigate this by using surrogate rewards and by initializing with domain knowledge (through causal insights and human input). Nonetheless, a limitation is that full autonomous optimization might be slow to reach its theoretical optimum, and we must monitor for any unintended consequences during learning. We should set expectations that Phase 4 is experimental and will require continuous tuning and oversight. Causal ML can help shorten the learning by pointing RL in the right direction, but causal inference itself can be tricky with observational data. The system might identify a "cause" that is not truly causal. Human validation remains in the loop to correct the course. In essence, while RL and causal AI promise adaptive improvement, they are **not magic wands** – they require robust data infrastructure and time to mature. We address this by piloting and gradually scaling the autonomous features, and by maintaining a parallel conventional approach as a fallback during the learning period (ensuring business continuity if the AI underperforms initially).

- **Complexity vs. Maintainability.** Combining multiple AI techniques (LLM, agents, RL, etc.) yields a sophisticated system, but also a complex one. There is a trade-off between performance and simplicity. We must ensure that the system remains **understandable and maintainable** by the technical team. One way we handle this is by modularizing each capability (as seen in phases) and using clear interfaces – for example, the output of Phase 3's engine is a defined set of insights that Phase 4's agent can consume, rather than an entangled black box. Documentation and monitoring will be crucial. We will implement dashboards for each component's performance (LLM output quality, agent actions, RL rewards, etc.), so we can pinpoint issues. Another best practice is to involve end-users and domain experts continuously: their feedback will often catch problems the metrics miss. For instance, if reps find the agent's plans impractical, that's a cue to adjust the logic even if the AI thought it was optimal. Agile iteration and perhaps model retraining or prompt refining will be part of ongoing maintenance. We acknowledge that an Agentic AI stack is cutting-edge – **talent and culture** are also factors. We will need to invest in training our team (data scientists, engineers, even TAMs) to work with these AI tools effectively. It's a socio-technical evolution, not just a software deployment.

- **Ethical and Regulatory Feasibility.** The plan assumes we can navigate regulatory guardrails around AI in healthcare commercial operations. We must be proactive in engaging compliance and legal teams at each phase. For example, Phase 1's explainable AI might raise questions: "Is the summary the AI provides considered promotional material that needs approval?" We should have a process to vet the types of outputs the LLM generates and possibly pre-approve templates or wordings. Similarly, by Phase 4, if the agent is largely autonomous in optimizing engagements, we need to ensure this aligns with FDA and industry guidelines (e.g. ensuring it doesn't suggest unapproved content). We also anticipate that guidelines for AI use might tighten (the **"AI Legislation"** mentioned in industry outlooks suggests stricter rules for explainability and consent [23]). Our roadmap bakes in explainability from the start, which should position us well to comply. We also maintain that the AI is assistive to reps, not replacing human decision-making – this is important for both legal and ethical acceptance. Physicians and patients likely would not respond well to a feeling of being "managed by an AI"; hence we keep the rep as the face and filter of all AI-driven suggestions, preserving the human touch in HCP interactions.

In conclusion, this phased roadmap is ambitious but grounded in the **current state of AI technology and business reality**. By starting with low-risk, high-reward enhancements (Phase 1's explainability) and incrementally adding autonomy and intelligence, we mitigate the chance of failure and build organizational confidence. Each phase delivers tangible benefits on its own, while laying the groundwork for the next. The ensemble of Agentic AI, LLMs, rules, and RL is our recipe for a solution that is both **innovative and responsible**. It acknowledges that no single AI technique can crack this complex problem alone – instead, we orchestrate them to complement each other's strengths [14]. The result, if executed well, will be a transformative capability for our oncology commercial team: an AI-driven system that *not only predicts the next best action, but understands it, explains it, orchestrates it, and learns from it*. This strategic advantage will help us drive growth, adapt swiftly to change, and ultimately better support HCPs and patients in their journey – all while maintaining the trust and compliance standards that define our industry.

Sources:

1. Russo, P. & Obayomi, I. (2025). *The power of data and machine learning to optimize HCP prescription volume*. PharmaLive – *On AI-driven switch targeting boosting new prescriptions* [1].
2. Soon, L. (2025). *Agentic Next Best Action*. Genesis: Human Experience in the Age of AI – *On combining LLMs, RL agents, and rules for NBA with explainability and compliance* [24] [3].
3. Malec, M. (2025). *Autonomous Agents: The Next Frontier in AI*. HatchWorks – *On how autonomous agents perceive, decide (using LLMs, RL, or both), act, and learn in a loop* [9] [14].
4. Mesinovic, M. *et al.* (2025). *Explainability in the age of large language models for healthcare*. Nature Communications Engineering – *On the critical need for transparent, explainable AI for clinicians and decision support* [2].
5. Genesis Blog (2025). *Agentic AI Challenges and Outlook – Noted challenges of latency, hallucinations (mitigated by retrieval), and the vision of NBA evolving into a self-optimizing, empathetic system* [6] [18].

---

[1] The power of data and machine learning to optimize HCP prescription volume - PharmaLive
https://www.pharmalive.com/the-power-of-data-and-machine-learning-to-optimize-hcp-prescription-volume/

[2] Explainability in the age of large language models for healthcare | Communications Engineering
https://www.nature.com/articles/s44172-025-00453-y?error=cookies_not_supported&code=940154cc-7f5d-463c-9d13-392573bada9e

[3] [4] [5] [6] [7] [8] [16] [17] [18] [19] [20] [21] [22] [23] [24] Agentic Next Best Action – Genesis: Human Experience in the Age of Artificial Intelligence
https://genesishumanexperience.com/2025/02/22/agentic-next-best-action/

[9] [10] [11] [12] [13] [14] [15] Autonomous Agents: The Next Frontier in AI | HatchWorks AI
https://hatchworks.com/blog/ai-agents/autonomous-agents/