

One agent, two teams

The agent. Medical affairs and commercial colleagues ask it data questions in plain language. It writes a query, runs it against the data warehouse, looks at the result, decides whether it needs another query, and drafts the answer with a chart. An analyst reviews the draft before it goes back. It works alone: it calls no other agent. It has run for three months with a pilot group of ten people at two companies.

The ask at both companies. Open it to every medical affairs and commercial colleague.

Score each team on each area. Tick the level 3 statements on your card that the team can show: none 1, some 2, all 3; all of level 3 plus some of level 5: 4. A statement that cannot apply to this design counts as shown.

Area	Team A	Team B	A	B
1 Business outcome	The head of medical analytics owns it. Target: routine requests answered in one day instead of five. Baseline measured before build; results reported monthly to leadership.	Owner and target named. Baseline measured before build.		
2 Risk appetite	A written rule, by type of action: it may read and draft; nothing leaves analytics without an analyst's approval. It applies to every agent, and the query tool rejects anything but reads.	The same written rule, enforced the same way.		
3 Workflow design and recovery	A documented rule hands an ambiguous question to an analyst. An analyst can hold an answer, but there is no agreed way to switch the agent off or recall an answer already sent.	Documented handoffs: it stops and passes to an analyst when the question is ambiguous or data is missing. A documented procedure lets the team switch it off and recall any draft; analysts answer by hand meanwhile.		
4 Human-in-the-lead	A named analyst, trained on the data marts and given time for every review, checks each answer with the query, the tables and the numbers in view, and has authority to stop the workflow. How often analysts correct answers is tracked monthly.	A named analyst, trained on the data marts and given time for every review, checks each answer with the query and tables in view, and has authority to stop the workflow.		
5 Data provenance	Every number links to its query, table and data version, automatically, including the version on the date asked. Only certified, documented data marts are reachable.	Every number links to its query, table and data version. Only approved, documented data marts are reachable.		
6 Agent identity and entitlements	It runs on the head of analytics' own login, which can read and change every table in the warehouse. The logs show her name. Nobody has listed the tools it uses, and code it writes runs directly on the warehouse server.	Its own service identity, read access to approved data marts only, every query logged as the agent's. Its tools are approved and listed; any code it writes runs in a sandbox.		
7 Evaluation and monitoring	Tested on past requests with known answers and pass thresholds agreed first, including trick questions. Production answers scored against them; any model or prompt change reruns the tests.	Tested on past requests with known answers and pass thresholds agreed first, including trick questions. Production answers scored against them.		
8 Unit economics and impact	Cost per accepted answer tracked, including review, corrections, retries and support, with spend and step limits, against the one-day target. Cost per answer decided which model it uses.	Cost per accepted answer tracked, including review, corrections, retries and support, with spend and step limits, against the one-day target.		

Ceiling (lowest of areas 2 to 7): Team A _____ | Team B _____

Which team would you open it up to, and why? _____