

# The Agency After Headcount

## The old agency model and the new agency model in life sciences content production

*Medware Group, September 2026. A working paper. Market figures are drawn from the Medware market baseline (02/09/2026) and carry the source and confidence ratings given there.*

### Summary

For fifty years the healthcare agency has sold the same thing: people. It bills hours, marks them up, and grows by adding heads. That model built the largest marketing networks in the world and it still dominates life sciences, where the top hundred North American healthcare agencies took in more than US\$11 billion in fees in 2023 and the two largest holding companies merged in 2025 with healthcare as their biggest client sector. It is also, in the specific case of regulated content production, coming to the end of its useful life.

The reason is not that clients have stopped needing content. They need more of it than ever: volume in the leading content platform doubled between 2019 and 2022. The reason is that the labour the agency sells is no longer the scarce input. A small senior team with AI production tools can deliver a review ready pharmaceutical asset with roughly one fifteenth of the human hours a conventional agency uses, and at a third of the price still keep a gross margin the agency cannot approach. When the scarce input changes, the pricing model built on the old one stops making sense, first for the client and then for the agency.

This paper describes the old model as it actually works, why it is straining, what the new model looks like in practice, what changes for the client and for the agency, and what does not change at all. It is written from the position of a company that has built the new model, and it tries to be honest about the limits of that position.

### 1. The old model, as it actually works

The healthcare agency sells time. Its rate card is a list of people by seniority, from junior staff at US\$100 to 150 an hour to principals at US\$300 to 500, and its estimate for any piece of work is a stack of those hours. A monthly retainer of US\$10,000 to 30,000, or in the enterprise tier the high six figures a year, buys a standing allocation of that time. The client is charged what the work cost the agency in salaries, multiplied by a markup that on an observed project was exactly two: US\$30,000 of agency cost billed at

US\$60,000. On that project, a two page flyer and an HTML infographic, the agency spent something like 240 hours of staff time.

This produces a particular shape of business. Revenue grows in proportion to headcount, so growth means hiring, and the profit of the enterprise comes from the gap between what junior staff cost and what they bill. The staffing pyramid, a few seniors over many juniors, is not an accident of culture; it is the margin. Holding company operating margins run at 15 to 20 per cent, and the healthcare networks inside them, at about US\$1.8 billion each for Omnicom and Interpublic in the first nine months of 2024, are the most profitable parts because pharmaceutical clients tolerate high rates and long engagements.

Regulated content adds a second feature. Every promotional piece must pass medical, legal and regulatory review before use, and the FDA received more than 150,000 such pieces in 2024. Review takes 16 to 21 days in the system and, by one vendor's measure, 50 to 60 days from concept to approval, with one to three rounds of revision. Each round is billable. The old model does not merely tolerate review friction; it earns from it. A third feature is the same: localisation, adaptation, and channel variants (an email, a slide, a banner from one approved claim) are all new hours, and content volume has grown 20 to 29 per cent a year on the leading platform's figures partly because each of those variants is produced from scratch.

It is important to be fair to this model. It carries strategic thinking, brand stewardship, creative judgement and long client relationships, and it absorbs the risk of a regulated environment on the client's behalf. Those are real services and clients have paid for them willingly. The point of this paper is narrower: the production of the content itself, which is most of the hours and most of the invoice, no longer needs the machinery that produces it.

## 2. Why it is straining

Four pressures have arrived together, and each on its own would have been survivable.

**The client has noticed the waste.** The most striking published figure in the market is not a cost but a proportion: about 80 per cent of approved pharmaceutical content is rarely or never used, according to field data across more than 600 million healthcare professional interactions. Clients are paying agency prices, and reviewer time, for material that does not reach a doctor. Once a commercial leader has seen that figure, the conversation with the agency changes from "what does this cost" to "why are we making it".

**The client has learned to insource.** Novo Nordisk built a modular content operation, grew from 5,000 to 31,000 assets a year, cut asset creation from 104 days to 23, and reported 60 per cent cost avoidance. Pfizer built an internal generative platform with the aim of tripling or quintupling content output. These are single cases and vendor reported, but they are the cases the rest of the industry studies, and they establish that a client can own the production function.

**The offshore hub has industrialised production.** Indegene, listed in India and earning more than 70 per cent of its revenue in North America, grew 18 per cent in US dollar terms in its 2026 financial year, delivers 83 per cent of its work offshore, generates about US\$75,000 of revenue per employee, and states openly that its strategy is to take share from agencies and CROs through AI driven centralisation. It has bought a regulatory writing firm and an Omnicom communications agency to do it. The hub is the old model with cheaper labour and better process; it competes on the agency's own terms and wins.

**The platform is moving into production.** Veeva, whose software runs review and content management for 47 of the top 50 biopharma companies, bought an agentic review company in June 2026 and launched a product it says will remove 70 per cent or more of manual review labour within five years. Adobe, Aprimo and EVERSANA make similar claims for their tools. The step where the agency billed most reliably, the review round, is being automated by the company that owns the review system.

Behind all four sits the expectation created by the consultancies. McKinsey has put a number on generative AI in pharma of US\$60 to 110 billion a year, with content creation costs down 30 to 50 per cent and review two to three times faster. Whether or not those figures prove right, every commercial leader in the industry has read them, and every agency proposal is now measured against them. An agency that arrives with the old rate card is arguing against a number its client already believes.

**The tell.** When an industry's incumbents all announce the same saving, the saving is no longer a differentiator; it is the new baseline. In 2026 the agency networks, the offshore hubs, the platforms and the consultancies all promise 30 to 70 per cent from AI. The question a client should ask is not who promises the most, but whose promise is on the invoice.

### 3. The new model

The new model starts from a different scarce input. If AI can do the drafting, the layout, the versioning, the reference checking and the first pass of compliance, then the scarce input is no longer hours; it is judgement applied at the right moments, and the knowledge of a brand and its rules encoded well enough that the tools can use it. Everything else follows from that.

**It sells outcomes, not time.** The unit of sale is a finished, review ready deliverable at a fixed price: a detail aid, an email set, a three minute mechanism of action animation, a congress content set. The price is not a stack of hours, and there is no rate card of people. The client buys the thing it wanted and knows what it will cost before work starts.

**It is priced against the market the client already knows.** Because the incumbents' prices are visible on the client's own invoices, the new model can price against them and guarantee the difference. In Medware's case the catalogue price is about a third of the typical billed market price, with a guarantee that it will be at least 50 per cent below whatever the client documents as its current cost. The guarantee is only credible because the cost structure underneath it is so different from the agency's.

**It starts with knowledge, not hours.** Each brand begins with an onboarding step that encodes its guidelines, its claims and the regulations that apply to it (in the United States, FDA OPDP rules, the PhRMA Code, state law, plus the company's and brand's own rules) into a knowledge bank that is used to produce every later project. This is the part of the work that is genuinely human and judgement heavy, and it is the part that compounds: the tenth asset for a brand is cheaper and safer than the first.

**It proves before it produces.** A proof of concept is delivered early in every project, so that misinterpretation is caught when it costs almost nothing. The old model catches it in review round two, at agency rates.

**It builds compliance in rather than billing for it.** Every deliverable carries its own review documentation: the claims matrix, annotated references, safety information placement and change log that reviewers need. The medically compliant draft goes to the client's medical writers and reviewers with the references attached, so their time is spent confirming rather than correcting. The new model does not replace them and should not want to; medical accountability stays where the regulator expects it.

**It is platform agnostic.** Output goes into whatever the client runs, Veeva, Adobe or otherwise. The new model competes with the agency, not with the platform, and it should stay that way, because the platform is where the client's audit trail lives.

**It runs on a small senior team.** On the observed flyer project, 15 human hours against an estimated 240 at the agency, a ratio of about 16 to 1, with a stated minimum of 10 to 1 across the catalogue. The team is producers, not a pyramid; there are no juniors to mark up because the work the juniors did is done by the tools.

## The two models side by side

|                                       | Old agency model                                                                                | New agency model                                                                            |
|---------------------------------------|-------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------|
| What is sold                          | Hours by seniority; retainers                                                                   | Finished deliverables at a fixed price                                                      |
| How growth happens                    | Hiring; headcount scales with revenue                                                           | More brands per producer; tooling scales with revenue                                       |
| Where margin comes from               | Markup on junior labour (about 2 times cost); 15 to 20% operating margin at the holding company | AI does the labour; gross margin above 80% at a third of the market price on modelled costs |
| Attitude to review rounds             | Billable                                                                                        | A cost to be designed out; two rounds included                                              |
| Attitude to variants and localisation | New hours                                                                                       | Reuse from the knowledge bank and modular library                                           |
| Who holds brand knowledge             | The account team, in people's heads                                                             | An encoded knowledge bank per brand, plus the producer                                      |
| Risk of misinterpretation             | Found in review                                                                                 | Found at proof of concept                                                                   |
| Price certainty                       | Estimate, revised with scope                                                                    | Fixed, guaranteed against the client's current cost                                         |

|                                | Old agency model                                | New agency model                                                 |
|--------------------------------|-------------------------------------------------|------------------------------------------------------------------|
| Relationship with the platform | Agency works inside the client's Veeva or Adobe | Same; delivers into the client's platform                        |
| Medical accountability         | Client's writers and reviewers                  | Unchanged                                                        |
| What it does not do            | Little; full service                            | Media, medical writing, strategy retainers, accredited education |

## 4. What changes for the client

The most visible change is the invoice, but the more important ones are operational. Fixed prices remove the estimate, the change order and the quarterly reconciliation; a brand manager can plan a year of content against a catalogue. Proof of concept early means that the brand team engages when engagement is cheap, which most teams prefer once they have tried it. Review teams receive fewer rounds and a complete documentation pack. Medical and regulatory colleagues notice this change first, so they should be in the room when a new model is evaluated because they are the ones who can veto it.

Procurement will ask the questions it always asks. If it is a third of the price, what is missing; who is accountable for medical content; how is the guarantee proven; can the vendor handle volume. The honest answers are that nothing is missing from the deliverable but the agency's headcount; that accountability sits exactly where it sits today; that the guarantee is written against the client's own last two invoices; and that capacity is a stated number of brands, not an unlimited promise. A new model that cannot answer those plainly is not ready.

There is also a change in what the client should measure. The old model was measured on cost per asset and on the agency's responsiveness. The new model should be measured on cost per used asset, on days to a review ready draft, on the first cycle approval rate (the market benchmark is 74 to 75 per cent), and on reuse. These are the numbers that expose the 80 per cent waste figure, and a client that adopts them will find that the price is the least of what changes.

## 5. What changes for the agency

The economics invert. Revenue per human hour on the observed project was about US\$670 at the introductory price and would be US\$1,330 at the catalogue price, against an agency principal's US\$300 to 500 and an offshore hub's roughly US\$40. Gross margin on modelled costs sits above 80 per cent at catalogue and above 70 per cent on the introductory price. Growth no longer requires hiring ahead of revenue, so the business can be profitable at a size at which a conventional agency is still subscale.

Two disciplines replace the old ones. The first is metering: because the margin depends on human hours and tooling cost per deliverable, both must be measured on every project, and the pricing model

revisited when they move. The second is capacity honesty: a small senior team has a ceiling in the number of brands it can hold in its head at once, and the right response is to state that ceiling and grow it deliberately, not to hire a pyramid to fill it.

The agency also gives things up. It does not buy media. It does not write medical content and should be clear that it does not. It does not sell strategy on retainer, though it may sell it by the project. It does not run accredited education. Those are large markets, and the new model's discipline is to leave them to others and to partner where a client needs the whole stack.

## 6. What does not change

The regulator does not change. Every piece still goes through medical, legal and regulatory review, still gets filed on Form 2253 at first use, and still faces an enforcement office that issued nearly a hundred letters in a single week of September 2025 after issuing five in the whole of the year before. A new model that treats compliance as a cost to be minimised rather than a product to be delivered will fail at the first audit, and deserve to.

The client's accountability does not change. Medical writers, medical reviewers and regulatory affairs remain the people who approve, and the new model's contribution is to make their job shorter, not to take it over. Brand strategy does not change; someone still has to decide what the brand should say, and an AI production model is a poor place to make that decision. Relationships do not change; pharmaceutical clients buy from people they trust, and a fixed price does not buy trust on its own.

And the platform does not change. Veeva, Adobe and the systems of record are where the client's content lives and will keep living. The new model succeeds inside them.

## 7. Objections and risks

**Quality.** The obvious objection is that AI produced content is worse. The honest answer is that it is worse without a knowledge bank, a senior producer and a proof of concept step. With them, it is at least as good. The way to settle the question is an introductory project at half price, which is why the new model should always offer one.

**The incumbents will catch up.** They will, in part. The platform will automate review; the hubs will add AI to their offshore teams; the holding companies will build internal tools as Pfizer and Publicis did. What they will find hard to do is abandon the pricing model, because the pyramid and the markup are the business. An agency that prices by the deliverable at a third of its own rate card is cannibalising itself, and holding companies are not built to do that voluntarily. The window is not permanent, but it is real, and it is measured in years rather than months.

**A race to the bottom.** If the new model is only a price, it will be undercut by the next new model. The defence is everything on the list above that is not a price: the knowledge bank that makes the tenth

asset better than the first, the documentation that makes reviewers prefer it, the metrics that show used assets rather than produced ones, and the guarantee that only a low cost structure can honour.

**Trust and data.** Pharmaceutical clients will ask where their brand knowledge lives, who can see it, and what happens to it when the engagement ends. A new model needs a clear, written answer, and in a regulated industry the answer should be conservative.

**The single data point.** This paper's figures for the new model rest on one observed project and a model that scales from it. That is stated plainly because it is true, and because the remedy is simple: meter the next five projects and publish the results.

## 8. The transition

Markets rarely switch models all at once. The likely path is that the new model wins the production layer first, one brand at a time, on introductory projects that prove quality and speed, while the old model keeps strategy, media and the relationship. Clients will run both for a while, and some agencies will try to become both. The hubs will compete hardest, because they are already priced on volume; the platforms will compete indirectly, by making review cheaper for everyone. Consolidation will continue, and the disclosed multiples of 2.5 to 3 times revenue for content and writing businesses in 2025 and 2026 suggest that new model companies with proven metrics will be bought rather than beaten.

For a company building the new model now, the practical conclusion is to be disciplined about three things: price by the deliverable and never by the hour; measure everything, especially the human hours and tooling cost that the margin depends on; and treat compliance and the client's reviewers as the product's first customers. The agency after headcount is not an agency without people. It is an agency where the people do only the work that needs them, and where the client pays for the outcome rather than the pyramid.

## Sources

Market figures: Medware market baseline, 02/09/2026 (MM+M Agency 100 2024 and 2025; Omnicom and Interpublic filings and Fierce Pharma, Dec 2024; Buzzbox Media and Healthcare Success pricing guides 2026; Veeva Pulse and benchmark reports 2020 to 2025; Veeva Copli release, 23/06/2026; Indegene FY2026 results and analyst meet, 2026; FDA OPDP Brief Summary, Jan 2025; Hall Render, Dec 2025; McKinsey, Jan 2024; Indegene Future of MLR Review, May 2026; Digiday on Pfizer Charlie, Feb 2024; Veeva Novo Nordisk case, May 2022; Certara and Indegene transaction disclosures, 2025 to 2026). Medware figures: observed flyer project (agency cost US\$30,000, billed US\$60,000, Medware US\$10,000, 15 human hours) and the Medware pricing model v1, assumptions marked in that model.