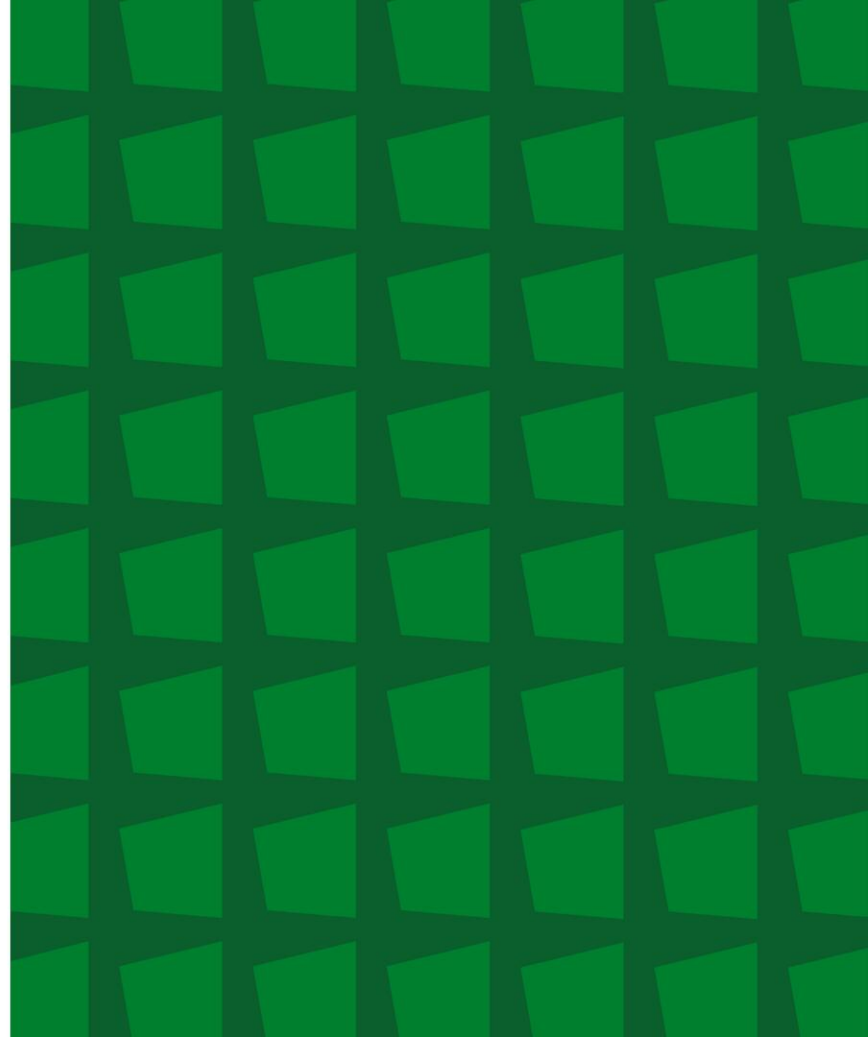


Build with AI, Run Without It

What AI can do today, and where hospitals should start

Falk Uebernickel · University of St. Gallen · St. Gallen, 28 October 2026

From insight
to impact.

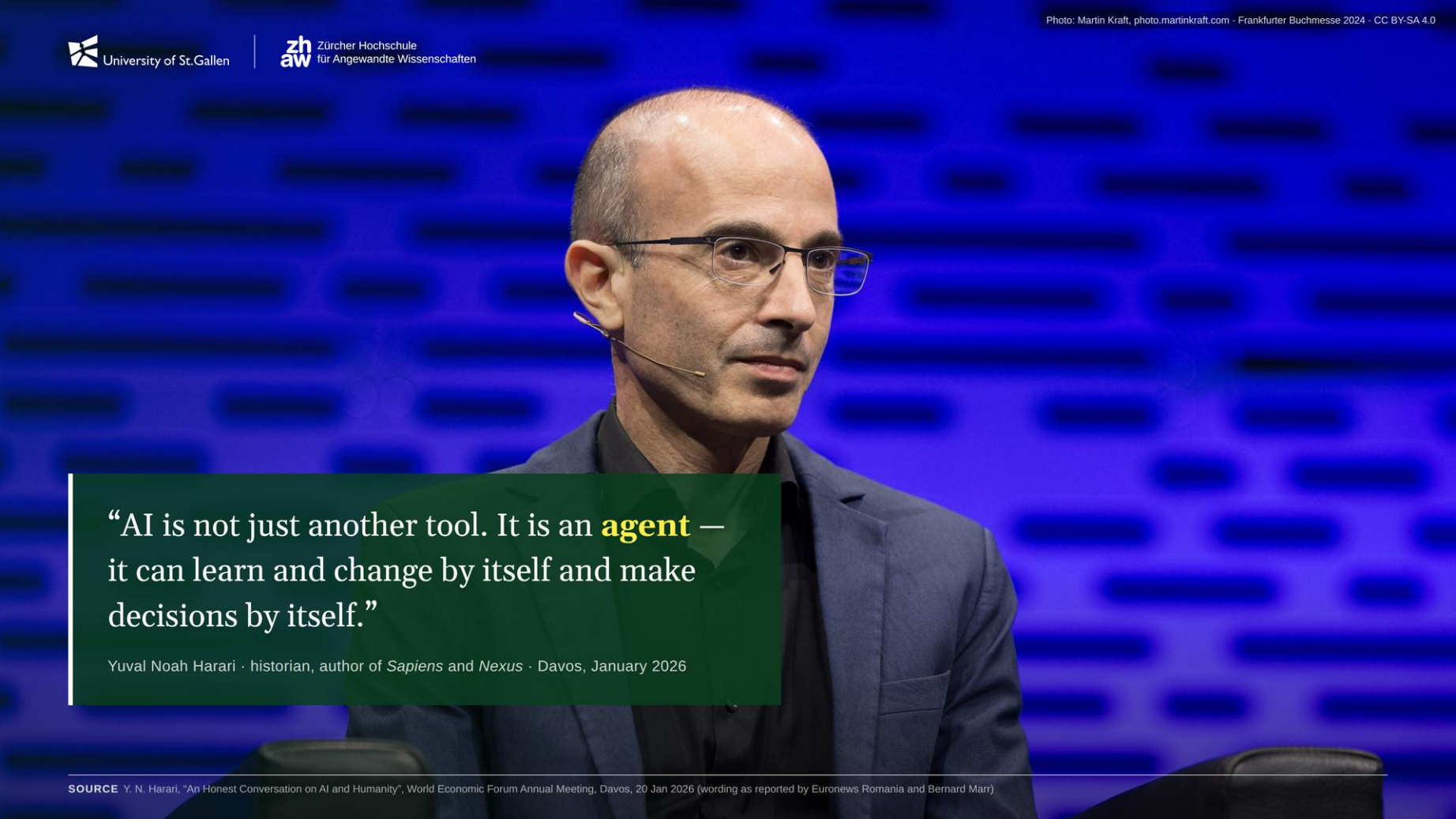


CIRCA 370 BCE · PLATO, PHAEDRUS

Every new technology was once feared

“This discovery of yours will create forgetfulness in the learners’ souls, because they will not use their memories; they will trust to the external written characters and not remember of themselves.”

— Socrates, in Plato's *Phaedrus*, 275a



“AI is not just another tool. It is an **agent** —
it can learn and change by itself and make
decisions by itself.”

Yuval Noah Harari · historian, author of *Sapiens* and *Nexus* · Davos, January 2026

01

PART 01

What AI can do today

How AI evolved, how it compares with people, what it costs and what agents can do.

Three eras in four years: AI moved from answering to acting, and the person from asking to approving

Three eras · One hospital example per era · Era boundaries: own framing

	Chat era · 2022/23 AI answers	Copilot era · 2024/25 AI assists inside tools	Agentic era · 2026 AI plans and acts
WHAT AI DOES	Answers in a chat window	Drafts inside the clinical system	Plans steps and acts across systems
IN A HOSPITAL	A chatbot answers patient questions 79% of expert ratings preferred the chatbot's answer <small>STUDY</small>	An ambient scribe drafts the visit note 63% of US hospitals on Epic use an ambient scribe (Jun 2025) <small>VIA PRESS</small>	An agent drafts the approval request -42% submission time at one US group <small>VENDOR-REPORTED</small>
THE PERSON	<small>HUMAN</small> Asks and decides	<small>HUMAN</small> Edits and signs	<small>HUMAN</small> Sets the task and approves
TASK LENGTH ¹	4 min GPT-4, Mar 2023	1 hour Claude 3.7 Sonnet, Feb 2025	12 hours Claude Opus 4.6, Feb 2026

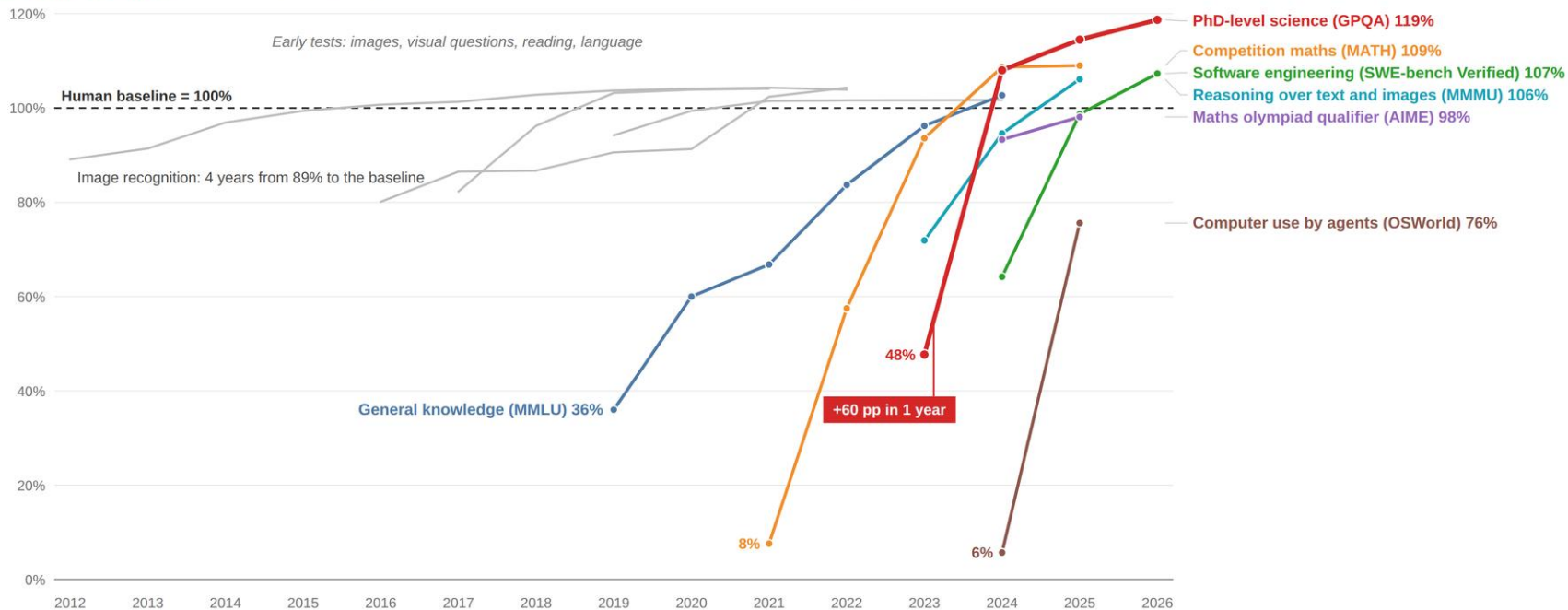
SOURCES ① Ayers et al., Comparing physician and AI chatbot responses to patient questions, JAMA Intern Med 2023;183(6):589–596 · ② METR, Time Horizon 1.1 benchmark results, accessed 24 Oct 2026

Further sources: TechTarget (Feb 2026, reporting an AJMC study) · Healthcare IT News, HIMSS26 (Mar 2026; Epic-reported, Penny at Summit Health, no baseline given) · Note: ¹ Longest software task the best model completes with 50% success (METR), log scale; not clinical work. Eras are own framing

AI now scores above the human baseline on 9 of 11 benchmarks, and newer tests fall faster: PhD-level science questions went from 48% to 108% in a single year

11 AI Index benchmarks · 2012–2026 · Grey: early tests, above the baseline since 2016–2021

Best AI model each year, % of the human baseline for the task (above 100% = better than people) · Benchmarks test well-defined tasks, not everyday work



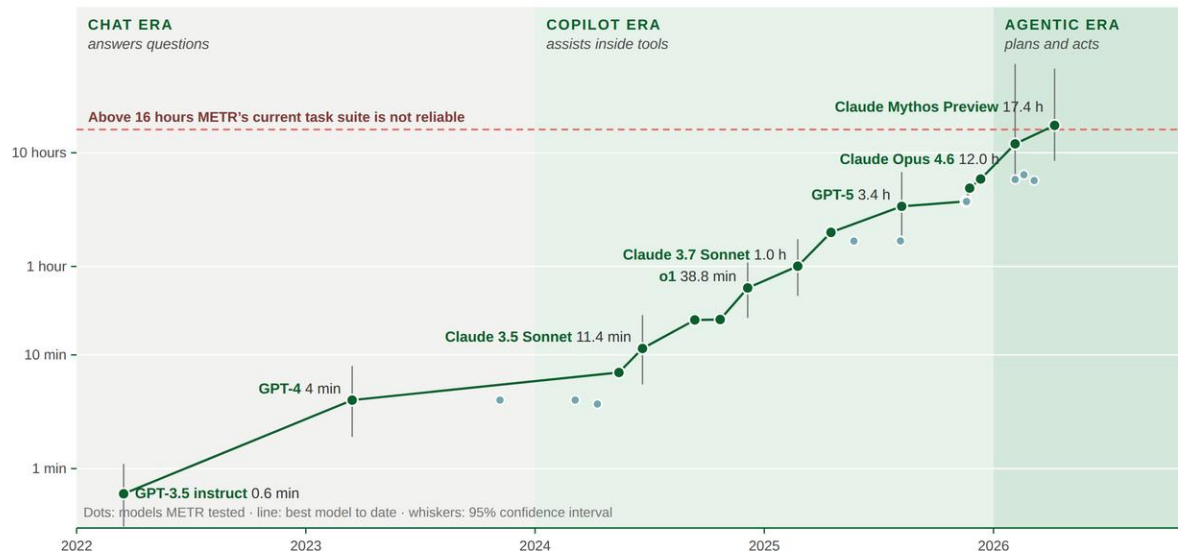
SOURCES ① Stanford HAI, AI Index Report 2026, Fig. 2.1.1 (Select AI Index technical performance benchmarks vs. human performance) · ② Epoch AI, Benchmarking Hub: best reported GPQA Diamond (96.4%, Sep 2026) and SWE-bench Verified (83.5%, Apr 2026)

Further sources: Stanford HAI, AI Index Report 2025, Fig. 2.1.33 · Note: Best model per year scaled to the human baseline of each test by the AI Index; values read from the published chart (±0.5 pp). 2026: Epoch scores scaled with the baseline implied by the AI Index (GPQA 81.2%, SWE-bench 77.8%). Latest values: SQuAD 2.0 2021; ImageNet, VQA 2022; SuperGLUE, MMLU 2024; MATH, MMMU, AIME, OSWorld 2025

In four years AI went from answering questions to completing multi-hour expert tasks on its own: the length of task a model can complete doubles about every four months

Task length over time · METR time horizon 1.1 · Eras: own framing

HOW LONG A TASK AI CAN COMPLETE ON ITS OWN METR 50% time horizon: length of software tasks, in human expert time, that a model completes in half of the attempts; log scale



129 days

doubling time of the task length since 2023 (95% CI 104–158); 188 days over 2019–2026

17.4 hours

Claude Mythos Preview (early), April 2026; 95% CI 8.5–55 hours, above the reliable range

~260×

longer tasks in three years: GPT-4 (Mar 2023) 4 minutes, Claude Mythos Preview (Apr 2026) 17.4 hours

SO WHAT FOR HOSPITALS

Raw capability is no longer the bottleneck for document, code and interface work; checking the output and deciding where the model runs are

THREE ERAS IN FOUR YEARS

Chat 2022/23 · Copilot 2024/25 · Agentic 2026: the best model went from minute-long to multi-hour tasks (in human expert time)

SOURCES ① METR, Time Horizon 1.1 benchmark results (benchmark_results_1_1.yaml), accessed 24 Oct 2026 · ② METR, Measuring AI ability to complete long tasks (2025)

Note: software and ML tasks only, not clinical work; the METR page has not been updated since 8 Sep 2026, so the newest models (e.g. GPT-6 Astra, Claude Opus 5.5, Claude Fable 5.1) are not measured; doubling time and CIs as published by METR; values above 16 hours are flagged by METR as unreliable; era boundaries are own framing

The same AI capability became 60 to almost 400 times cheaper within two years

Cost · Price of the cheapest model that reaches a fixed capability level · Blended price, 3 input : 1 output tokens

USD PER MILLION TOKENS, CHEAPEST MODEL AT A FIXED CAPABILITY LEVEL *log scale; each dot = a new cheapest model*



The top tier stays expensive

USD per million tokens, blended 3 : 1

GPT-4 list price, Mar 2023

37.50

Top tier today GPT-6 Astra, Claude Fable 5.1, Oct 2026

20.00

GPT-4 level cheapest model (Llama 3.1 8B), Jul 2024

0.10

1.9×

cheaper: top-tier list price, Mar 2023 to Oct 2026

375×

cheaper: GPT-4 capability, in 16 months to Jul 2024

SOURCES ① Stanford HAI, AI Index Report 2025, Fig. 1.3.22 (inference price at constant performance), with Epoch AI data (LLM inference price trends) · ② Vendor pricing pages of OpenAI and Anthropic (vendor), current and archived via Internet Archive Wayback Machine, read 24 Oct 2026

Note: Levels: GPT-3.5 on MMLU, GPT-4 on HumanEval, GPT-4o on GPQA Diamond. The AI Index 2025 data end in early 2025; the AI Index 2026 contains no inference-price data. Where curves meet, the same model was cheapest at both levels. List prices are not adjusted for tokenizer differences. Fold changes and blended prices are own calc.

Agents already prepare prior-authorization requests at US health systems, on early evidence

One task, end to end · Prior authorization; in Switzerland: cost approval

Epic  **Summit Health**

-42%

less time to submit a medication prior authorization when Epic's agent drafts it

VENDOR-REPORTED

Epic

25

health systems each built their own AI agent for a complete task at an Epic workshop, clinical leads with IT and Epic staff

VENDOR-REPORTED

Stanford

91–98%

of test tasks an AI agent completed correctly in a simulated patient record (98% with memory)

BENCHMARK

KLAS Research

0.5%

of AI applications in US health care were agents at the end of 2025

SURVEY



13 hours

a week per US physician and their staff go to prior authorization. In Switzerland the matching step is the **cost approval** by the insurer

SOURCES ① Healthcare IT News, Epic highlights AI systems' success metrics at HIMSS26 (Mar 2026; Epic-reported) · ② Epic Systems, At Agent Factory Build-a-Thon, health systems create their own AI agents (15 Sep 2026, vendor)

Further sources: MedAgentBench v2 (Stanford), Pac Symp Biocomput 2026 · KLAS via healthsystemCIO (31 Dec 2025) · AMA prior authorization survey (n = 1,000), 2025 · Note: -42%: Epic's Penny at Summit Health, no baseline or sample given; no independent or Swiss evaluation yet. Logos identify the source or the organisation reported

02

PART 02

Where hospitals stand

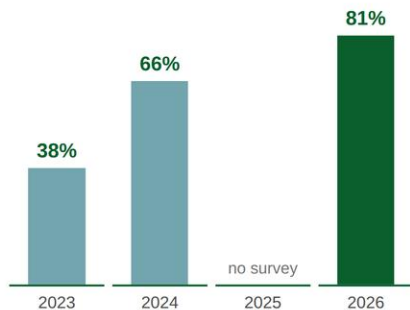
Adoption in the US and Switzerland, department by department.

In the US, AI use became the norm within three years: four in five physicians now use it at work

United States as lead market · 2016–2026 · Three independent measures

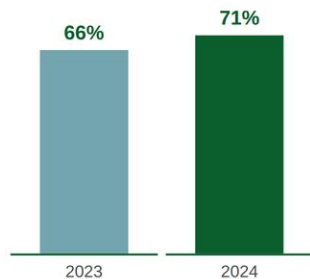
PHYSICIANS

US physicians using AI in their work (AMA survey)



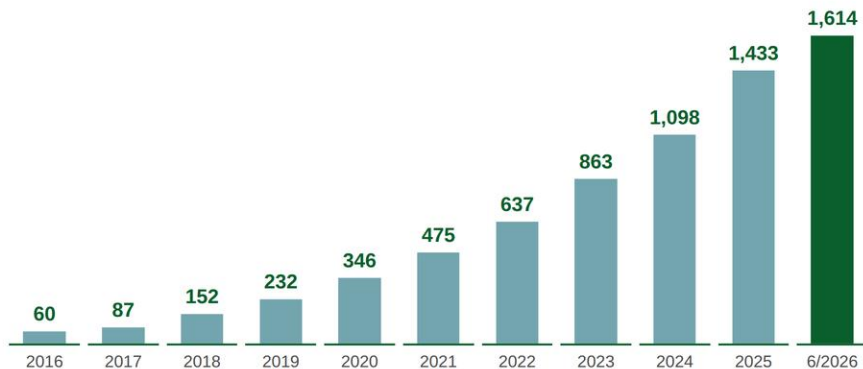
HOSPITALS

US non-federal acute-care hospitals with predictive AI in the EHR



AUTHORISED DEVICES

AI-enabled medical devices authorised by the FDA, cumulative



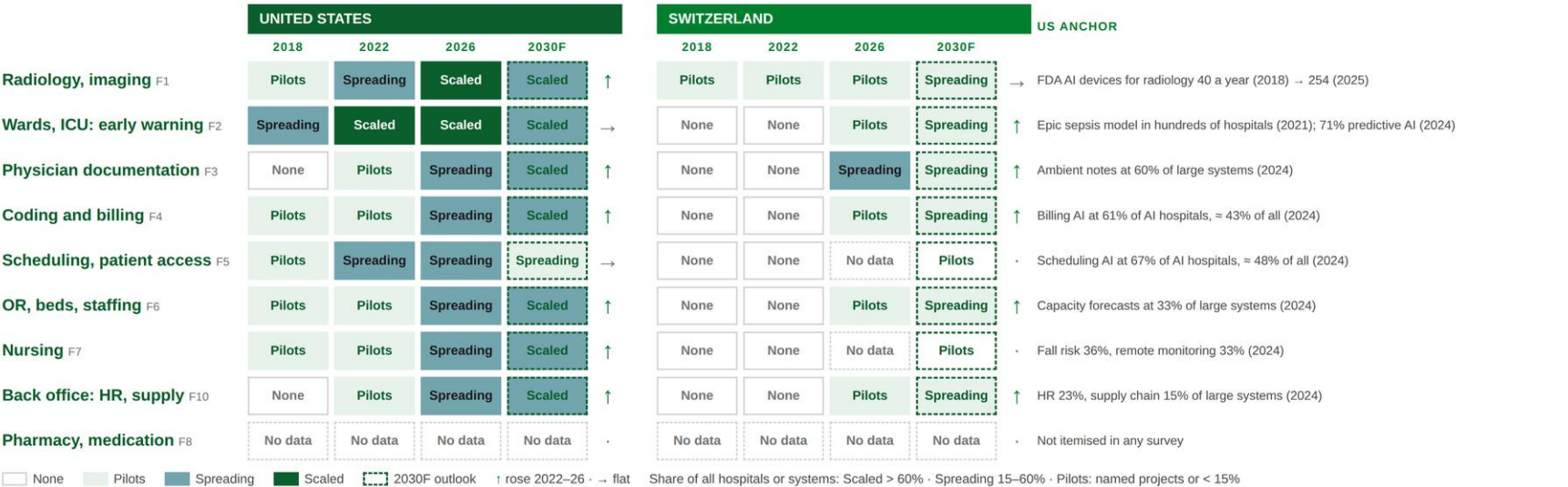
No AMA survey wave in 2025 · ASTP data for 2025 not yet published · FDA: authorisations to 29 June 2026 (list status 4 Sep 2026); about 76% of devices are in radiology

SOURCES ① American Medical Association, Physician Survey on Augmented Intelligence (2023, 2024, 2026) · ② ASTP/ONC (HHS), Data Brief 80, Hospital Trends in the Use of Predictive AI 2023–2024, based on AHA IT Supplement

Further sources: (3) JAMA Network Open, generative AI in US hospital EHRs, based on the AHA IT Supplement 2024 · U.S. FDA, List of AI-Enabled Medical Devices (own count, list status 4 Sep 2026) · Note: The FDA list is not comprehensive and is revised retroactively

In the US, AI use rose in 6 of 9 hospital departments between 2022 and 2026; in Switzerland only documentation has moved beyond pilots

AI adoption by department · US vs Switzerland, 2018 to 2026, outlook 2030 · Levels: own judgement on the cited evidence



SO WHAT

Overall, Switzerland in 2026 is about where the US was in 2022; in Switzerland documentation leads, the rest is pilots

Outlook 2030 assumes Switzerland follows the US with a lag; pharmacy and medication AI is not measured anywhere

SOURCES ① ASTP/ONC (HHS), Data Brief 80, Hospital Trends in the Use of Predictive AI 2023–2024 (2025) · ② Poon et al., JAMIA 2025, Scottsdale Institute survey of 43 US health systems (Fall 2024)

Further sources: FDA AI-enabled device list (4 Sep 2026); Wong et al., JAMA Intern Med 2021 (Epic sepsis model); Swiss hospital and company releases 2021–2026 (see H2); KPMG Clarity on Healthcare 2026 · Note: Levels 2018/2022 reconstructed by the author from the cited evidence; 2030F by rule: US one step up where it rose 2022–26, Switzerland at most one step up and not beyond the US 2026 level

Inside Swiss hospitals AI is still at pilot stage: 88% of CIOs and CTOs say at most a fifth of AI pilots reach production

Inside the hospital · four of seven dimensions · Switzerland only, public sources, status October 2026 · Named examples are reported facts, not ratings of any hospital

DIMENSION	KEY FIGURES	REPORTED FACTS AND EXAMPLES	STATUS · OWN ASSESSMENT
Expectations <i>What do leaders expect?</i>	88% 65% of hospital CIOs and CTOs say at most a fifth of their AI pilots reach production (n = 8) ^C of hospital CFOs expect AI to lift EBIT by 2% or less; none expects more than 4% (n = 24) ^C	48% of CFOs plan to invest only 0.1–0.5% of revenue in AI until 2027 ^C . Biggest hurdle: data protection and regulation (88% of CIOs/CTOs), budget: none	Pilot Spreading Scaled Many pilots, few in production
Use cases <i>Where is AI in use?</i>	114 min a day go to documentation for Swiss acute-care physicians, up from 85 min in 2011 (2025)	Kinderspital Zürich: report time of residents “almost halved”, moving to routine ^H . USZ simulation trial: consultation time –18%, documentation quality lower · LUKS: AI duty-roster drafts, planning effort about –2% ^V · USB: LLMs tested in five departments ^S · none of six large emergency departments uses AI for triage (May 2026) ^S	Pilot Spreading Scaled Documentation leads; few results measured
Organisation <i>Who owns AI?</i>	— No representative Swiss data on AI roles or committees exist, and no hospital chief AI officer appears in public sources	Inselspital: first international site in the ACR ARCH-AI programme for radiology AI governance (2026) ^S · USB: own organisational structure for ethics and safety of AI (2024) ^S · Study of 28 Swiss health companies (hospitals excluded): 10 had no AI ethics in their processes (2026)	Pilot Spreading Scaled Single bodies, mostly radiology
Core systems <i>Is the digital base ready?</i>	CHF 182.5m is the one-off cost of the Insel Gruppe's Epic programme: CHF 83m contract, go-live 2024, more than 50 systems replaced	USZ Epic: award CHF 94.8m (2025), estimate now about CHF 190m ^S · CHUV with 11 Vaud hospitals: about CHF 89m plus about CHF 10m a year ^S · EPD: 95.6% of hospitals connected, 123,599 records opened (Sep 2025)	Pilot Spreading Scaled Core systems being renewed

Evidence tags: ^C consultancy survey, small n · ^V vendor or company claim · ^S secondary source or trade press · ^H hospital-reported, method not published · no tag: official statistic, study or the institution's own document

SOURCES ① KPMG Switzerland, Clarity on Healthcare 2026 (hospital CFOs n = 24, CIOs/CTOs n = 8; consultancy) · ② gfs.bern for FMH, Befragung zum ärztlichen Arbeitsumfeld 2025

Further sources: Insel Gruppe media release 11.02.2026 · Netzwoche (USZ Epic 07.10.2026; CHUV 21.11.2025; USB 09.12.2024) · Kinderspital Zürich news 01.06.2026 · J Clin Anesth 2026 (USZ trial) · Microsoft Schweiz 2024 (LUKS, vendor) · Nau.ch 21.05.2026 ·

AuntMinnie Europe 2026 · AI and Ethics 2026 · BAG EPD fact sheet 2025 · Note: Status ratings are own assessment; selection, not a ranking

03

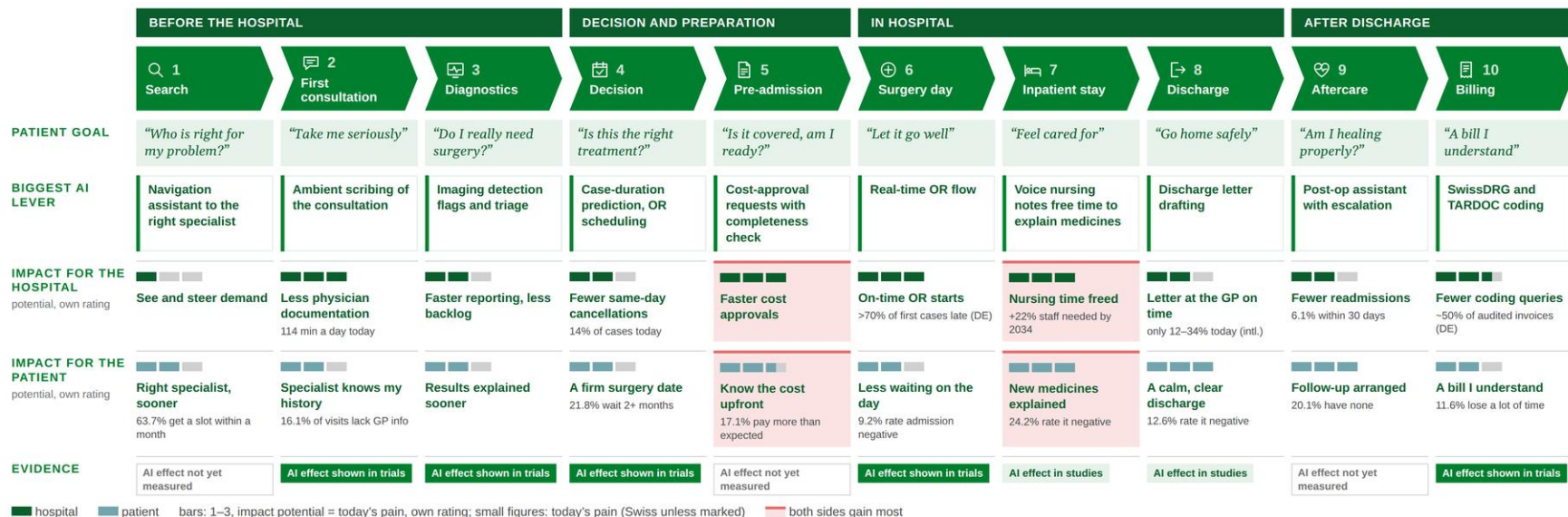
PART 03

Where AI meets the patient

The patient journey, AI by patient segment (one example), what patients want versus what providers offer, and where it hurts most.

AI can help in every phase; hospital and patient gain most at once at the cost approval before admission and on the ward

Elective patient, Swiss private hospital · 10 phases, one AI lever each · Impact for hospital and patient: own rating from today's pain points



WHAT THIS MEANS FOR HOSPITALS

Start where hospital and patient both gain: the cost approval before admission and care on the ward

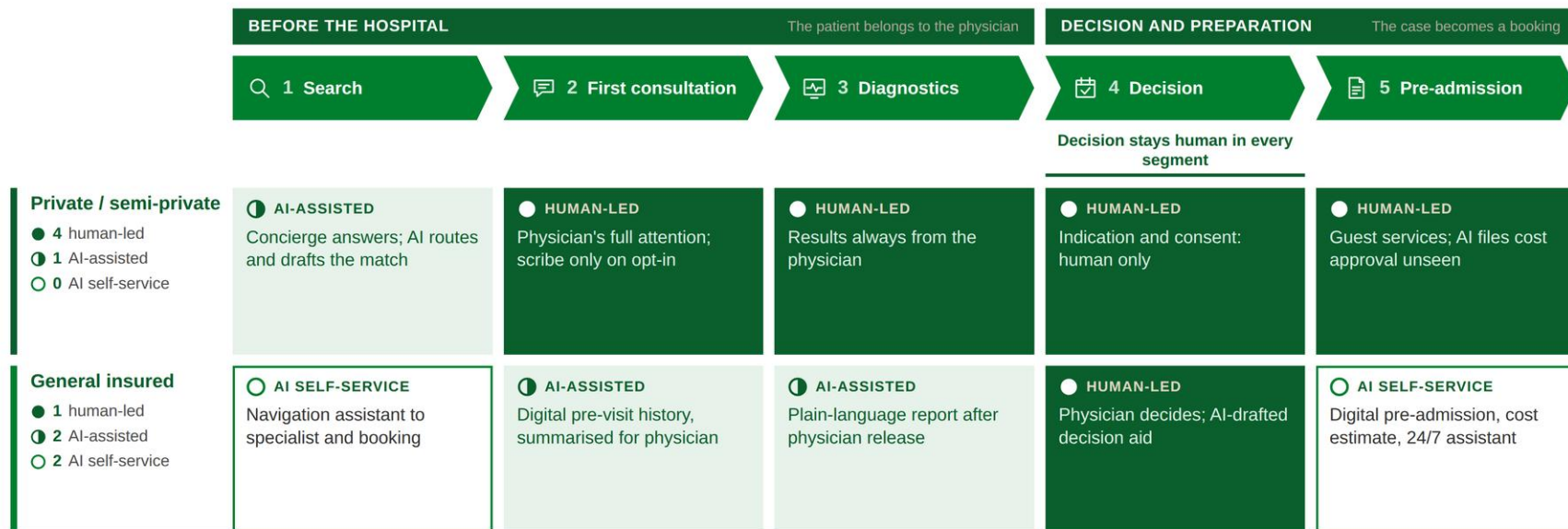
The levers with the strongest evidence (documentation, OR, coding) help the hospital more than the patient; aftercare helps the patient but is not yet measured

SOURCES ① ANQ, Nationale Patientenbefragung Akutomatik 2025 (PREMs); Obsan 10/2023 (Commonwealth Fund IHP 2023) · ② gfs.bern for FMH 2025; Obsan 2026 nursing forecast

Further sources: Anaesthesia 2026 (cancellations, meta-analysis); Dtsch Arztebl Int 2013 (first-case delays, DE); Kripalani et al., JAMA 2007; ANQ readmissions 2023; MD Bund 2026 (DE); evidence of AI effects: see the opportunity map · Note: Journey and levers: own synthesis, not data of any specific hospital; impact potential is today's pain intensity 1–3 per side (own rating)

Before admission premium patients meet a person in 4 of 5 phases; general patients book and prepare by self-service in 2

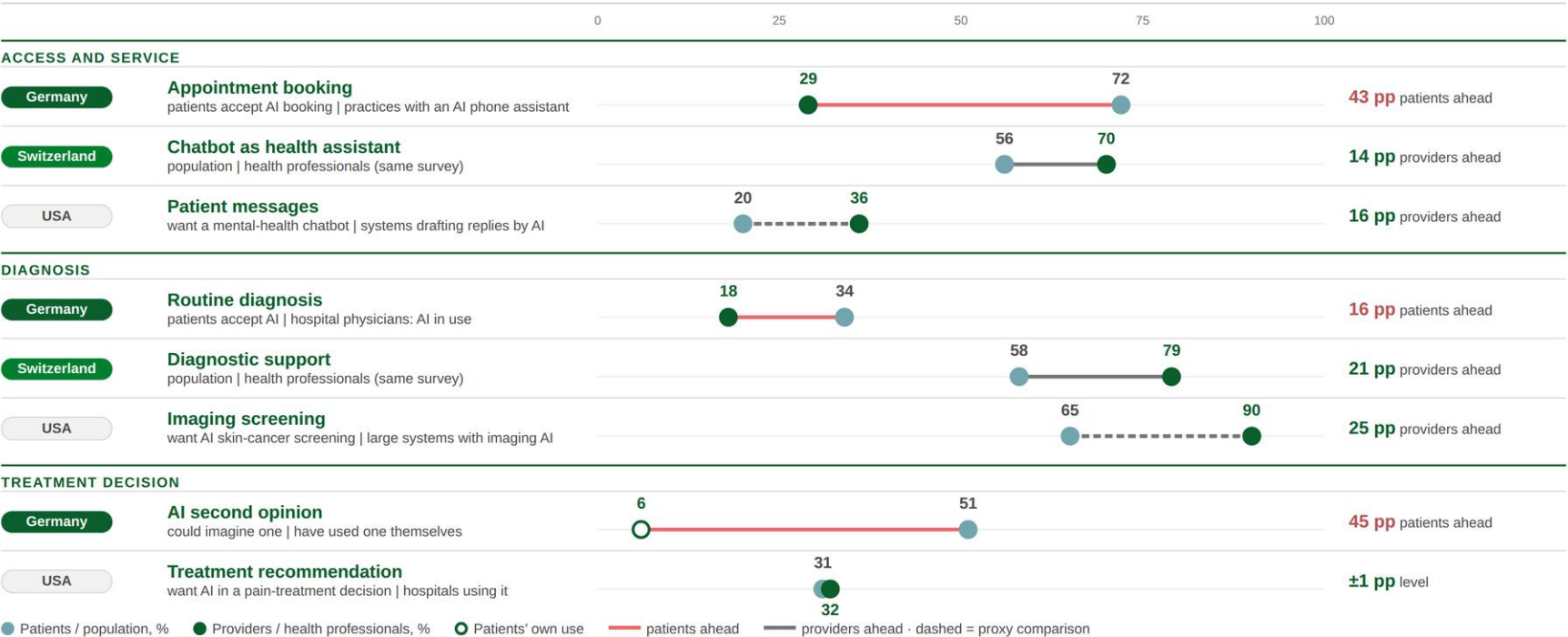
Front-stage AI dosage, phases 1–5, own design · ● **Human-led** AI invisible, backstage ① **AI-assisted** output released by a named person ○ **AI self-service** one tap to a person



SOURCES ① Own dosage design on the elective patient journey of a Swiss private hospital; evidence per cell from own research (Oct 2026) · ② Longoni, Bonezzi, Morewedge (2019), Resistance to medical AI, J Consumer Research 46(4)
 Further sources: Information Systems Research 2025 (service robots in luxury); Doctolib Digital Health Report 2026 (vendor); Pew Research Center 2023 (n=11,004); ANQ PREMs acute care 2025; Obsan 10/2023 (Commonwealth Fund IHP 2023) · Note: Partner practices own phases 2–4; dosage there needs their agreement

German patients want AI booking far more often than practices offer it, while in diagnosis US and Swiss providers are ahead of patients

Patients vs providers · Germany · Switzerland · USA · Rows compare different surveys: read the direction, not exact gaps



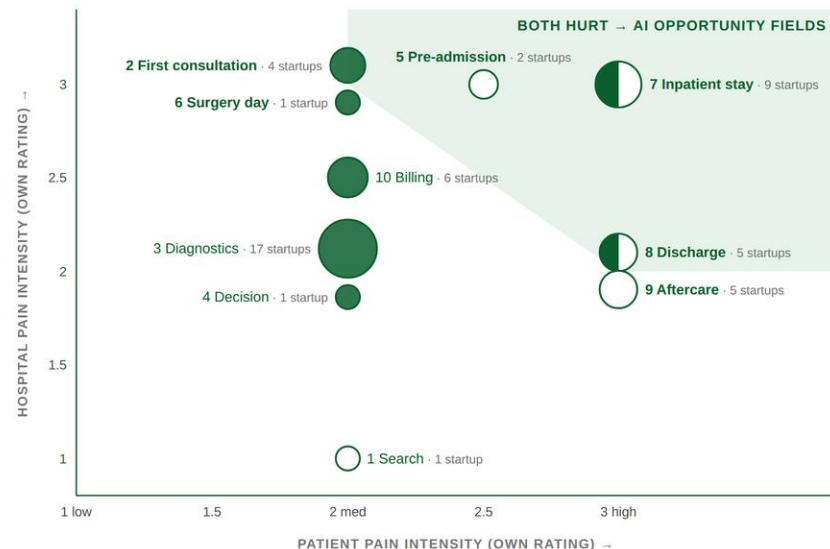
SOURCES 1 Doctolib Digital Health Report 2026 (vendor); Medizininio 2026 (vendor); Bitkom 2024/2025 · 2 gfs.bern, Swiss eHealth Barometer 2026 · 3 Pew Research Center (Dec 2022, n = 11,004); ASTP/ONC Data Brief 80 (2025); Poon et al., JAMIA 2025; Eliciting Insights 2026 (vendor)

Note: US hospitals recommending treatment by AI is an own derivation (share with predictive AI × share using it for the purpose); Swiss rows compare what the population and health professionals can imagine, no Swiss deployment share exists

Both sides hurt most on the ward and at pre-admission, where AI benefit is least proven; startups crowd diagnostics (17 of 39), a phase of medium pain

10 journey phases, 2 perspectives · 6 opportunity fields · AI benefit ● RCT ● other study ○ none

ONE DOT PER PHASE · SIZE = STARTUPS ACTIVE IN THE PHASE · FILL = AI BENEFIT OF THE MAIN LEVER



TOP 5 OPPORTUNITY FIELDS

ranked by pain sum, then measured AI benefit, then startups

- Inpatient stay** pain 3 + 3 = 6
 Voice nursing notes free time to explain new medicines
 8 of 9 startups in the phase match this lever · ● other or indirect study · L2 nursing, 0.3–1.4% of cost
FIRST STEP Pilot voice notes on one ward; track the ANQ medicines score
- Pre-admission** pain 2.5 + 3 = 5.5
 Cost-approval request with completeness check, out-of-pocket estimate
 1 of 2 startups in the phase match this lever · ○ no measured benefit · L3 admin and back office, 0.6–1.9% of cost
FIRST STEP Measure days to approval and refusal rate; pilot with one insurer
- First consultation** pain 2 + 3 = 5
 Ambient scribing and pre-visit summary
 4 of 4 startups in the phase match this lever · ● measured in RCT · L1 physician documentation, 0.1–0.4% of cost
FIRST STEP Offer a scribe to partner physicians in one practice centre
- Surgery day** pain 2 + 3 = 5
 Case-duration prediction and real-time OR flow
 0 of 1 startups in the phase match this lever · ● measured in RCT · L9 OR capacity: upside, not a cost pool
FIRST STEP Replay one clinic's OR timestamps against a duration model
- Discharge** pain 3 + 2 = 5
 Discharge letter drafting and medication explanation
 4 of 5 startups in the phase match this lever · ● other or indirect study · L1 physician documentation, 0.1–0.4% of cost
FIRST STEP Letter drafted on the ward, sent to the GP on discharge day

Next: **Aftercare** (sum 5, no measured benefit, 5 of 5 startups match this lever)

SOURCES ① ANQ, Nationale Patientenbefragung Akutsumatik 2025 (PREMs, n = 46,923, 186 hospitals) · ② Obsan Bericht 10/2023, Commonwealth Fund International Health Policy Survey 2023, Switzerland

Further sources: gfs.bern for FMH 2025 (documentation minutes); ANQ readmissions 2023 (release 2025); Obsan care-staff projection 2024–2034; DE first-case delay study 2013 (22 hospitals); Anaesthesia 2026 meta-analysis (same-day cancellations); Kripalani et al., JAMA 2007; MD Bund (DE audits); CMS HCAHPS 2024 (US, same pattern); evidence: published AI benefit studies; value pools: own lever model (% of total hospital cost) · Note: intensity 1–3 per phase is own judgement from the published evidence; startups: own screening of 39 CH/DE startups (a startup can serve several phases, so the dot counts add up to 51; billing includes startups with a billing feature; the list counts only startups that match the lever's function); quadrants: Verification Matrix v2 (else own analogy)

04

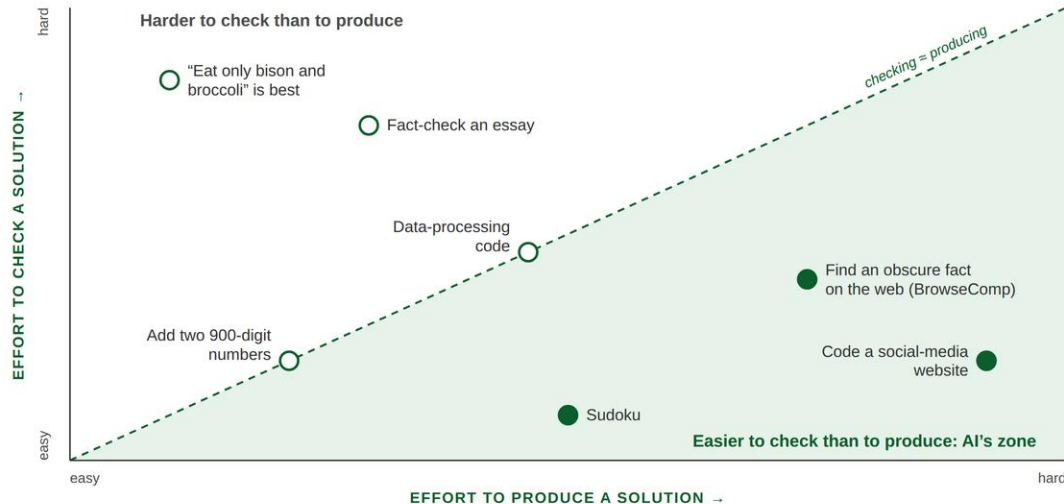
PART 04

How humans stay in charge

Verifier's law, the Verification Matrix, what the hospital buys versus designs, and how roles change.

AI learns fastest where a solution is cheap to check: tasks that are hard to produce but easy to verify fall first

Verifier's law · Asymmetry of verification · Positions qualitative, after Wei (2025)



Verifier's law

"The ease of training AI to solve a task is proportional to how verifiable the task is."

Jason Wei (2025)

What makes a task easy to check

- 1 Objective truth: everyone agrees what a good solution is
- 2 Fast: one solution is checked in seconds
- 3 Scalable: many solutions are checked at once
- 4 Low noise: the check tracks true quality
- 5 Continuous: solutions can be ranked, not only passed or failed

SOURCE ① Wei, J. (2025), Asymmetry of verification and verifier's law, blog post, jasonwei.net

Note: Redrawn after Wei's examples; positions show the direction of the asymmetry, not measurements

The Verification Matrix gives each task its own form of human oversight: 7 of 12 operational and mixed AI use cases need Sign-Off

Verification Matrix with 12 operational and mixed hospital AI use cases · Placement and design control: own assessment

VERIFICATION EFFORT: how hard is it to check? ↑

C Proof Trail

OVERSIGHT AT MILESTONES

Claims stay assumptions

- The output looks simple; whether it is right shows only later
- People check at milestones

OPERATIONAL AND MIXED USE CASES · 1

- Admission and bed forecast, weeks ahead

D Co-Creation

CONTINUOUS OVERSIGHT

Build together, release by gate

- Hard to produce and hard to check
- Nothing leaves without explicit approval

OPERATIONAL AND MIXED USE CASES · 1

- Discharge letters M

A Autopilot

OVERSIGHT AFTER THE FACT

Act and undo

- AI acts on its own
- People watch the log and correct mistakes

OPERATIONAL AND MIXED USE CASES · 3

- Scheduling and admin chatbot
- OR, bed and staffing forecasts, next day
- MRI no-show prediction

B Sign-Off

OVERSIGHT PER OUTPUT

Draft and approve

- AI drafts at volume; each item is quick to check
- An expert approves every output

OPERATIONAL AND MIXED USE CASES · 7

- ☐ Ambient scribing
- ☒ Coding (SwissDRG, TARDOC)
- ☒ Replies to patient messages M
- ☒ Duty-roster drafting
- ☒ Chart and pre-visit summaries M
- ☒ Cost-approval requests to insurers
- ☐ Evidence look-up with citations M

TASK COMPLEXITY: expertise and context it takes → ☐ Embedded ☒ Configured ☒ Built by the hospital M = mixed process

OUTSIDE THE MATRIX

Human only

Decisions a person keeps by choice, not because AI could not support them

Breaking bad news

Informed consent

End-of-life decisions

Final indication and treatment decision

Coercive measures in psychiatry

ICU triage under scarcity

Credentialling of clinicians

SO WHAT

Almost every task can be AI-supported; the question is not whether, but which form of human oversight fits the task

Where a hospital buys AI, it verifies through procurement and local proof; where it designs AI, it can design the checks

Who verifies each layer, and how · Three levels of design control × four verification layers · Layer model: own proposal

	Embedded	Configured	Built
	Bought inside a device or product <i>Vendor controls logic and change</i> e.g. Imaging flags, ECG reading, mammography reader	Vendor platform, set up by the hospital <i>Hospital sets parameters, not the core model</i> e.g. Sepsis warning, coding, scheduling, rosters	Designed by the hospital <i>Hospital controls logic and change</i> e.g. Discharge letters, cost-approval requests, tumour-board prep
L1 Product Does the AI do what it claims, in general?	MANUFACTURER; REGULATOR IF A DEVICE The hospital secures in the contract: version pinning, notice, rollback, pre-authorised updates (PCCP)	VENDOR, CORE MODEL Hospital sets parameters, not the core model	HOSPITAL Own test set, versioned prompts and models; regulation route first: MDR Art. 5(5), MepV
L2 Site Does it work here, on our patients and data?	HOSPITAL Local validation on own data before signing, silent mode	HOSPITAL Thresholds, triggers and recipients set and tested on local data	HOSPITAL Release gate: tested on own data before go-live
L3 Use Is each output, and its use, right?	HOSPITAL, CLINICIANS Checkable outputs as award criterion: sources, confidence, explanation; automation-bias tests	HOSPITAL Templates, sources and who signs what	HOSPITAL Cheap checks designed in: source links, structured fields, rules in the output
L4 Outcome Does the benefit materialise and last?	SHARED Performance and incident data back from the vendor	HOSPITAL, WITH VENDOR DATA Drift and the response to an alert	HOSPITAL Sampling plans: reviewers re-check a share of outputs over time
	EXAMPLE Sepsis models: one external check found AUC 0.63, another setting –18.7% mortality; certification did not predict which		EXAMPLE Discharge letters: sentence-level source links and medication checks move them from Co-Creation to Sign-Off
	HOSPITAL the hospital verifies this layer itself	VENDOR manufacturer or vendor verifies	SHARED both, with vendor data

SO WHAT

A hospital can buy the model, but not the proof that it works on its own wards

Buying moves one of four layers to the vendor; for the other three the hospital decides at purchase what it can check later

8 of 16 hospital roles have grown or emerged since about 2015 – in nursing, informatics, data, AI oversight and security – while transcription and front-desk work shrink

Role change in the hospital · 16 roles · c. 2015 · today · c. 2030

	DIRECTION ¹	C. 2015	TODAY	C. 2030 ²	EVIDENCE ANCHOR
CLINICAL					
Registered nurses	Growing	●	●	● F	Swiss demand for nursing staff +22% by 2034; graduates cover 40–55% (Obsan 2026)
Clinical informaticists, CMIO	Growing	●	●	● F	US certified clinical informaticians 1,105 (2016) → 2,496 (ABPM, 2024) (JAMIA; ABPM)
Pharmacists	Stable	●	●	● F	US employment +5% projected, 2025–35; 23% of hours automatable by 2030 (BLS; MGI)
Physicians	Changing	●	●	● F	Swiss hospital doctors: documentation 101 → 114 min/day, 2015–25 (gfs.bern/FMH)
Radiologists and pathologists	Changing	●	●	● F	US radiologists using AI 30% (2020) → 68% (2024) (ACR; Clin Imaging)
ADMINISTRATION & OPERATIONS					
Bed and capacity management	Growing	●	●	● F	25 of 31 surveyed US systems run a command centre; NHS study: no flow effect (JCJQPS 2023; NIHR)
Medical coders, records specialists	Changing	●	●	● F	US jobs +8% to 2035, but BLS flags AI coding; AI in billing 36% → 61% (BLS; ASTP)
Medical secretaries, admin assistants	Changing	●	●	● F	US projection cut from +21% (2014–24) to +5% (2025–35) (BLS)
Revenue cycle, prior authorisation	Changing	●	●	● F	63% of US providers use AI in revenue cycle; AI cut prior-auth drafting 42% (HFMA; Epic)
Medical transcriptionists	Shrinking	●	●	● F	US jobs 70,000 (2014) → 42,000 (2025) → 40,200 (2035, projected) (BLS)
Patient access, scheduling, call centre	Shrinking	●	●	● F	US hospitals using AI for scheduling 51% → 67% (2023–24); receptionists –2% (ASTP; BLS)
DATA, AI & GOVERNANCE					
Health information technologists	Growing	●	●	● F	US health information technologists +16% projected, 2025–35 (BLS 2025–35)
Data scientists in health	Growing	●	●	● F	US data scientists +35% projected, 2025–35 (all industries) (BLS 2025–35)
Cybersecurity in hospitals	Growing	●	●	● F	US health data breaches (500+ records) 270 (2015) → 741 (2024) (HHS OCR via HIPAA Journal)
AI governance / chief AI officer	Emerging	○	●	● F	US health systems with an AI governance structure 40% → 70%, 2024–25 (HFMA)
Clinical AI validation, monitoring	Emerging	○	●	● F	82% of US hospitals with predictive AI test it locally; 58% monitor after go-live (ASTP 2024)
Dot size = schematic weight of the role, derived from the direction label ● dashed = 2030 outlook F = 2030 outlook rests on a published forecast or projection; otherwise judgement					

¹ Direction labels are the author's classification from the sources, not a measured headcount trend; anchors are single indicators, mostly US, that illustrate the direction. Swiss hospital staff (FTE) grew 2015–24 in all groups, administration included. ² Outlook to 2030: BLS projections 2025–35, Obsan 2026, McKinsey Global Institute, WEF Future of Jobs 2025 where marked F.

SOURCES ① US Bureau of Labor Statistics, Occupational Outlook Handbook, projections 2014–24 and 2025–35 · ② ASTP/ONC, hospital AI use 2023–24 (AHA IT Supplement)
Further sources: gfs.bern/FMH 2025; Obsan 2026 (via ARTISET); ACR; JAMIA/ABPM; McKinsey Global Institute; HFMA 2024–25; Epic (vendor); Jt Comm J Qual Patient Saf 2023; NIHR; HHS OCR via HIPAA Journal; WEF Future of Jobs 2025

05

PART 05

Where hospitals should start

What hospitals build and what vendors deliver, where the startups are, and three decisions.

Hospitals can build in operational and mixed processes now, while treatment AI arrives embedded in vendor products

What hospitals build and what vendors deliver · Process type × horizon · Horizon: mainstream use in Swiss and German private hospital groups

	NOW 2026–27	NEXT 2028–29	LATER 2030+	
Hospital builds AI is the builder; much of the result runs without a model	Operational No personal data, or employee data	<ul style="list-style-type: none">■ Interfaces and adapters runs without a model■ Roster rules and staffing patterns runs without a model■ Cost-approval rules runs without a model	<ul style="list-style-type: none">■ Agents for operational tasks, read-only first model runs live■ KIS migration and code mapping runs without a model■ Discharge letters model runs live	–
	Mixed Patient data, used for administration			■ Cross-system clinical agents† model runs live
	Treatment Patient data in care decisions; often a medical device	Rarely built: a medical purpose brings device duties (MDR, MepV); the in-house exemption is narrow		
	Verify For everything, built or bought	<ul style="list-style-type: none">■ AI inventory incl. AI inside vendor products■ Release gate for every AI tool	<ul style="list-style-type: none">■ Local validation before EU duties (2028)■ Evaluation and monitoring in use■ Privacy tech, federated validation†	■ Synthetic data for training†
Vendors deliver Embedded in products or configured on platforms	Operational No personal data, or employee data	<ul style="list-style-type: none">□ AI cyber defence	<ul style="list-style-type: none">■ Operations digital twin, command centre■ Logistics and pharmacy robotics†	–
	Mixed Patient data, used for administration	<ul style="list-style-type: none">□ Ambient notes	<ul style="list-style-type: none">□ Ambient orders, codes and letters in the KIS†	–
	Treatment Patient data in care decisions; often a medical device	<ul style="list-style-type: none">□ Edge AI in devices□ Point solutions, e.g. fall sensors, delirium scores■ Imaging AI via one platform	<ul style="list-style-type: none">□ Generative imaging reports†■ Remote monitoring after surgery†	<ul style="list-style-type: none">□ Multimodal foundation models□ Autonomous narrow diagnostics†□ Surgical robot autonomy†
		■ Built by the hospital, with AI as the builder ■ Configured on a vendor platform □ Embedded in a bought product † horizon: own judgement		
Over time: design freedom is largest in operational and mixed processes and it is closing; what a hospital has not specified by about 2028 tends to arrive as a vendor default (own assessment)				

■ Cross-system clinical agents[†]

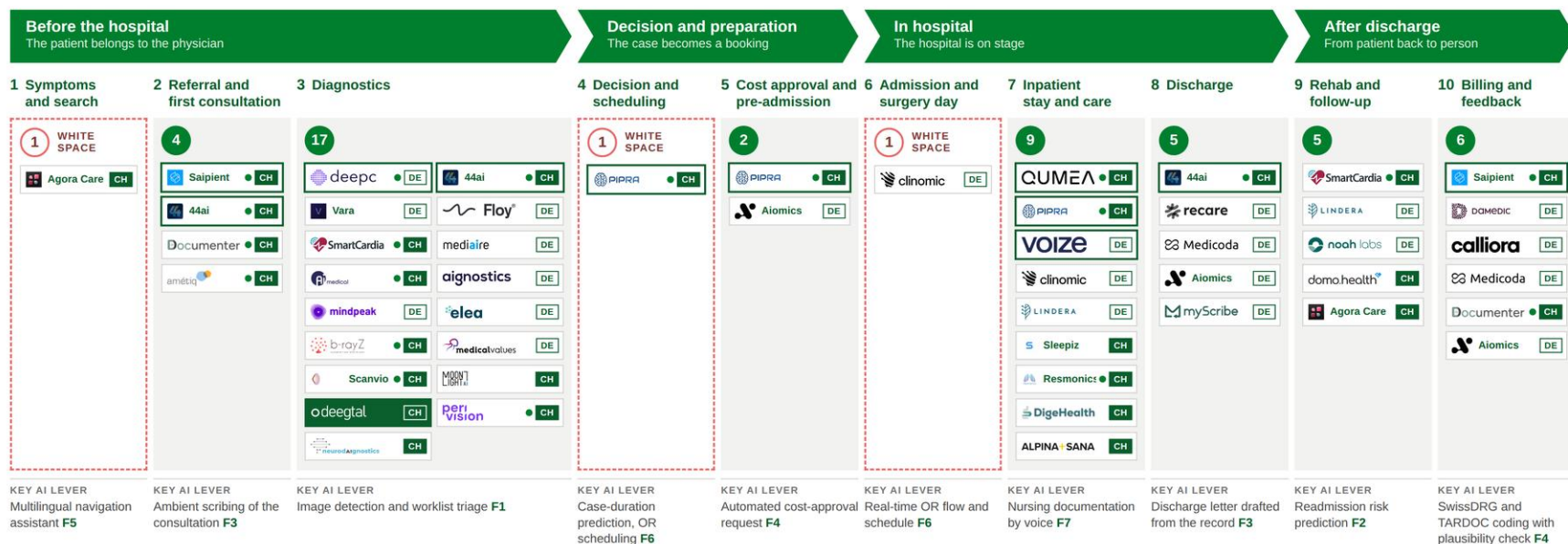
model runs live

SOURCES 1 Deloitte, Agentic AI in health care (Feb 2026); Gartner, 2025 agentic AI forecast; WHO, guidance on large multimodal models (2024/25); KLAS (Dec 2025); Epic, AI Charting (Feb 2026) · 2 Swiss DSG, StGB Art. 321, OR Art. 328b, ArGV 3 Art. 26; EU MDR Art. 2(1), 5(5) with MDCG 2023-1; Swiss MepV; Regulation (EU) 2026/1744

Further sources: Adams et al., JMIR 2025; SAPinsider 2024; UKE 2025; ASTP/AHA IT Supplement 2024; QUMEA, PIPRA; Dodsworth et al., Age and Ageing 2024 · Note: Process type as in exhibit V4; design control, horizons of built items and the re-sorting into mixed processes: own assessment

Startups crowd into diagnostics (17) and the ward (9); scheduling and cost approval stay almost empty, and the four startups built for DRG billing are all German

Startups along the elective patient journey · 39 startups, mapped to the phases they serve · CH Switzerland · DE Germany · ● Swiss hospital reference □ Top-6 candidate



Startups serving several phases appear in each (10 of 39). Outside the journey: ORamaVR (staff training). Key AI lever: one lever per phase from the journey; F-codes: hospital functions.

SOURCES ① Own startup database: 39 CH and DE hospital-AI startups, vendor websites, press releases, startupticker.ch, EU-Startups, tech.eu, startbase, Oct 2026 · ② Own generalised journey of an elective patient at a Swiss private hospital group, 2026; phase mapping: own

Note: Own assessment based on vendor-stated information (websites, press releases, startupticker, EU-Startups), status 24 Oct 2026

Three decisions for hospital leaders: build where the hospital owns the process, buy with proof, govern both kinds of AI

Conclusions · General conclusions for hospitals · Own assessment

1

Build where the hospital controls the process

Interfaces, cost-approval rules, pre-admission checks, migrations: AI builds, tests verify, the result runs without a model

2

Buy clinical AI with proof

Ask the checklist questions before signing; validate on your own data in silent mode; keep change control over updates

3

Run two governance regimes

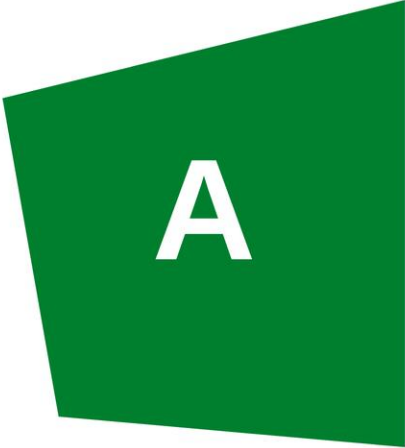
A release gate for what AI builds, lifelong monitoring for models that run in care; name builders, verifiers and model owners

THE LINE TO REMEMBER

AI is the architect of the automation, not the decision maker

SOURCES ① Own synthesis, Oct 2026 · ② Duke-Margolis Institute for Health Policy, AI governance in health systems (white paper, 2024)

Note: General observations for hospitals, not the plan of any specific hospital



A

PART A

Appendix

Agenda and further exhibits

Agenda

Five parts · From what AI can do today to where hospitals should start

- | | | |
|-----------|-------------------------------------|---|
| 01 | What AI can do today | AI went from answering questions to working for hours on its own, at a fraction of the cost |
| 02 | Where hospitals stand | The US has made AI routine across departments; Swiss hospitals are at the pilot stage |
| 03 | Where AI meets the patient | AI candidates run along the whole journey; premium and general patients want different help |
| 04 | How humans stay in charge | Almost every task can be AI-supported; the difference is who checks, and who designs |
| 05 | Where hospitals should start | Treatment AI arrives embedded; build what the hospital controls and buy the rest with proof |

Value estimates overshoot on speed: Germany realised 4% of its stated potential within four years, and two-thirds of Swiss hospital CFOs expect at most +2% EBIT

Promise vs reality · Germany and Switzerland · * own calculation from the sources' figures

1 Germany: promised vs realised

McKinsey value pools, digitalisation incl. AI, EUR bn a year

BY 2022

Potential stated in 2018

EUR 34bn

4%

Realised

EUR 1.4bn

of the 2018 potential
realised*

BY 2026

Realised a year

EUR 11.5bn

13%

Still open

EUR 80bn

of the total realised*

2 Swiss hospital CFOs expect little

Expected EBIT effect of AI (n = 24)

65%

expect AI to lift EBIT by 2% or less

≤ +2%: 65%

+2–4%: 35%*

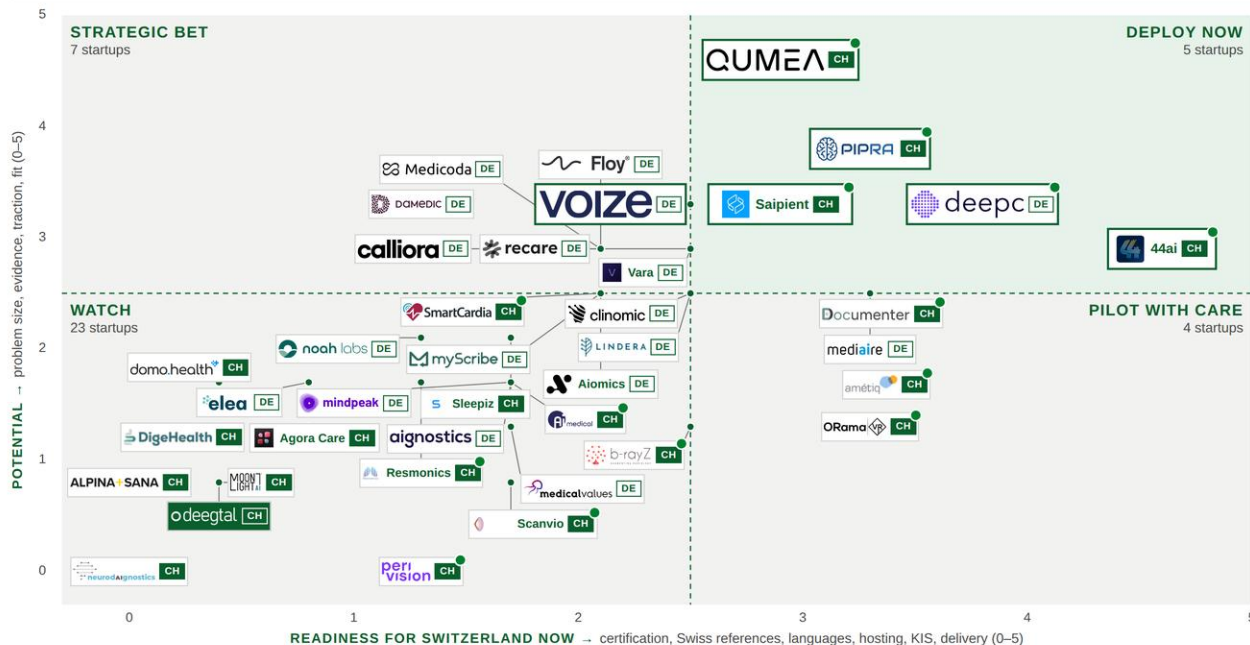
None expects more than +4%

SOURCES ① McKinsey & Company Germany, Die 42-Mrd.-EUR-Chance, May 2022, and Das 80-Mrd.-EUR-Update, Oct 2026 · ② KPMG Switzerland, Clarity on Healthcare 2026 (n = 24 hospital CFOs)

Note: McKinsey values are gross, system-wide and cover digitalisation including AI; 4% = 1.4 of 34; 13% = 11.5 of 11.5 + 80

Only five of 39 startups are both ready for Switzerland and aimed at a large problem

Readiness × potential · own scoring · 39 startups · CH Switzerland DE Germany ● Swiss hospital reference □ Top-6 candidate



STARTUPS PER QUADRANT

Deploy now	5
Strategic bet	7
Pilot with care	4
Watch	23

Deploy now: QUMEA, PIPRA, deepc, Saipient, 44ai

SENSITIVITY

If unknowns scored 1 instead of 0, six of the seven strategic bets (voize, Damedic, Vara, Floy, Recare, Medica) would move to Deploy now; Calliora would not

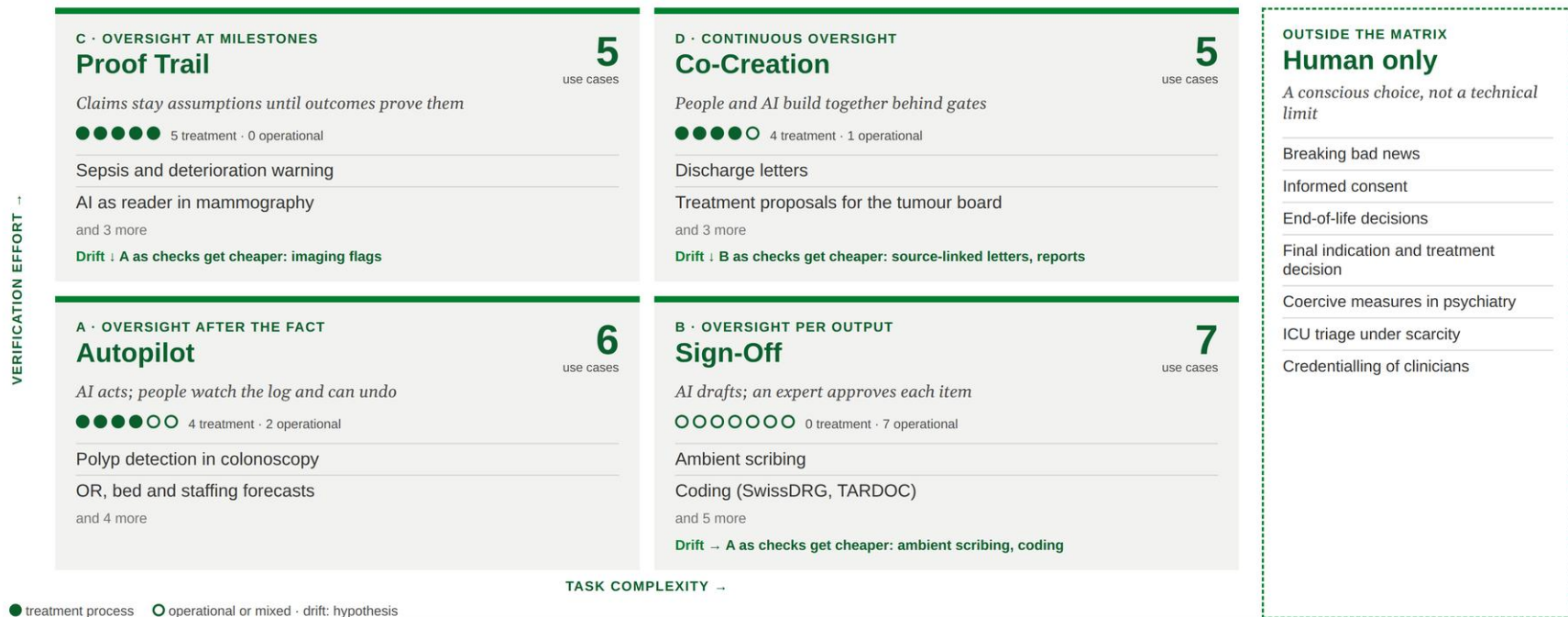
Due diligence decides: most gaps are unpublished hosting, integration and MDR class, not proven weaknesses

SOURCES ① Own startup database: 39 CH and DE hospital-AI startups, vendor websites, press releases, startupticker.ch, EU-Startups, tech.eu, startbase, Oct 2026 - ② Own scoring of readiness for Switzerland and potential for a Swiss private hospital group, six sub-criteria each, Oct 2026

Note: Own assessment based on vendor-stated information (websites, press releases, startupticker, EU-Startups), status 24 Oct 2026; scores are not measurements

Almost every hospital task can be AI-supported; what differs is how humans stay in charge

23 hospital AI use cases in four oversight models · Positions: own assessment, 5 binary criteria per axis (after Hötter 2026)



SOURCES ① Ueberrnickel, F. (2026), Every task can be AI-supported: the question is how humans stay in charge, working note, University of St.Gallen · ② Hötter, J. (2026), master's thesis, Hasso Plattner Institute

Note: Use-case positions are the author's assessment for hospitals, not validated with clinicians

Hospitals govern AI in five ways, from a central committee to an in-house institute; the committee is the most common, the others are single cases

Five governance models · One example house each · Public examples, no endorsement

1 Central committee



One board reviews every tool at fixed gates

A COMMITTEE DECIDES

EXAMPLE



Duke Health

52 algorithms under one oversight committee, checked from design to after go-live (2022)

Heidelberg University Hospital: AI governance board of about 10 functions

2 Executive AI role



A chief AI officer near the top runs a central team

ONE EXECUTIVE DECIDES

EXAMPLE



Mayo Clinic

Mayo: chief AI implementation officer reports to the CEO; 128 tools in clinical use

Also Cleveland Clinic: chief AI officer since 2024

3 Federated council



Sites and service lines decide under shared rules

EACH LINE DECIDES

EXAMPLE



Providence

Service-line governance across 50+ hospitals under a generative-AI leadership council

Decision rights still being chartered (2026)

4 Platform institute



An in-house AI institute builds and runs the tools

THE INSTITUTE DECIDES

EXAMPLE



Universitätsmedizin Essen

IKIM institute plus a smart-hospital platform; about EUR 14m public funding to 2027

No separate AI committee found

5 External standard



An outside programme sets the review standard

OUTSIDE BODY SETS THE BAR

EXAMPLE



INSELSPITAL
UNIVERSITÄTSPITAL BERNE
HÔPITAL UNIVERSITAIRE DE BERNE

First site outside the US in the ACR ARCH-AI programme for radiology AI (2026)

Covers radiology only

66% of US hospitals that use predictive AI have an AI committee (2024). The other models are single cases; no study compares which works best

SOURCES ① ASTP/ONC, Hospital trends in the use, evaluation and governance of predictive AI, 2023–2024 (Data Brief 80) · ② Duke-Margolis Institute for Health Policy, AI governance in health systems (white paper, 2024)

Further sources: Duke AI Health (ABCDS, JAMIA 2022) · Universitätsklinikum Heidelberg, KI-Governance-Board · Healthcare Innovation (Mayo Clinic) · Healthcare Finance News (Cleveland Clinic) · Healthcare IT News (Providence, secondary) · Land NRW and IKIM (Essen) · AuntMinnie Europe (InselSpital) ·

Note: Models are own grouping of published cases; figures as published by the institutions

AI builds the hospital tool; the tested tool runs without a model in the decision path

Concept · Two places for AI: design time and run time · Own framing

DESIGN TIME · AI BUILDS

1 Humans specify

the rule, the interface, the guideline

2 AI builds

code, mappings, rules, test cases

3 Humans and tests verify

the release gate decides, not the model

RUN TIME · NO MODEL

4 The artefact runs

same input, same output; no model inside

WHERE IT APPLIES IN A HOSPITAL

- **Lab results into the EHR:** AI builds the HL7/FHIR adapter; an integration-engine channel runs it
- **Cost approval requests:** insurer criteria become rule checks; staff send the request
- **System migration:** AI writes mapping and ETL scripts; reconciliation tests check them

THE RULE

- **No probabilistic model in the decision path**, which is not the same as “no AI at run time”
- **Deterministic means reproducible and testable**, not automatically correct

Where the input is speech, images or free text, a model may extract at the edge; a deterministic check verifies its output

SOURCE ① Own framing (Oct 2026); EU AI Act, Recital 12 (rules defined solely by people are not AI systems)

Further sources: MDCG 2019-11 Rev.1 and Swissmedic, Merkblatt Medizinprodukte-Software · Note: Hospital examples are illustrative, not measured; deterministic does not mean unregulated, as medical-device law (MDR, MepV) qualifies software by intended purpose; whether AI-drafted, human-verified rules count as defined solely by people is our reading

On numbers and codes AI adds no edge but cost and variance; it earns its place where it has to read speech, images or text

When AI should run, and when a rule is enough · Systematic reviews, trials, benchmarks

NUMBERS AND CODES: A FORMULA OR RULE DOES AS WELL

±0

AI predicts no better

than a classic statistical formula, across 145 comparisons in published studies¹

15 pts

AI is not repeatable

Same question, same settings, ten runs: the share of correct answers moved by up to 15 points **PREPRINT**

~200×

Checking costs more than the AI

A physician checking one AI-written note costs CHF 2.8–3.9; the AI call costs USD 0.016²

SPEECH, IMAGES, FREE TEXT: AI IS NEEDED

–44%

Images: less reading work

AI as one of two readers in breast screening cut radiologists' workload (randomised trial, Sweden)

0.92

Free text: earlier warning

Reading nurses' notes, the AI spotted patients who deteriorate better: score 0.92, against 0.75 without the notes (UK emergency department; 0–1)³

1.00

Text plus rule: no errors

AI reads the report, a rule checks the result: fully correct, against 0.63 for the AI alone (Bern; score 0–1)³

SOURCES ① Christodoulou et al., J Clin Epidemiol 2019 (logistic regression vs machine learning, systematic review) · ② BMJ Health & Care Informatics 2024 (free-text triage notes vs NEWS, UK ED)

Further sources: Atil et al., arXiv 2408.04667 (preprint) · Anthropic pricing, Sonnet 5.5 list price (vendor) · JMIR Med Inform 2026 (scribe editing time, Sweden) · MASAI RCT, Lancet Digit Health 2025 · Commun Med 2025 (Bern, 206 reports) · Note: ¹ AI = machine learning; formula = logistic regression; difference in AUC 0.00 in low-bias studies. ² Own arithmetic: median 93 s editing per note at one Swedish provider, at an assumed CHF 110–150 per hour; model call: 4,000 input + 800 output tokens at USD 2 / USD 10 per million tokens; UK and Bern results are single-site and retrospective; logistic regression is a transparent statistical model, not a hand-written rule. ³ Average precision (UK) and F1 (Bern), both 0 to 1, higher is better

Building with AI meets a real staffing gap, but its gains are not yet measured in hospitals

Evidence on the design-time side · Surveys and trials · Vendor-affiliated figures marked

ADOPTION OUTRUNS TRUST

84%

of developers use or plan to use AI tools (2025; 70% in 2023)

48%

trust AI only where its output is easy to verify (2026)

HOSPITAL IT LACKS BUILDERS

76%

of German hospitals cannot fill open IT posts

64%

of hospitals with IT vacancies cite little room to develop own software

MEASURED GAINS ARE MIXED

+56%

faster on a small new task (Copilot RCT, 95 freelancers) **VENDOR AUTHORS**

+21%

faster on a complex task (Google RCT, 96 engineers) **GOOGLE AUTHORS**

-19%

slower for experts in their own large codebases (METR RCT, 16 developers)

—

no measured AI-built HL7/FHIR interface in any hospital yet

SOURCES ① Stack Overflow Developer Survey 2024, 2025 and 2026 (AI sections) · ② BDO and Deutsches Krankenhausinstitut, Personalnotstand im Krankenhaus – Quo vadis? (2023, n = 111)

Further sources: Peng et al. 2023 (GitHub, vendor-affiliated) · Paradis et al., ICSE-SEIP 2025 (Google) · METR 2025 · Note: No Swiss data published on AI coding tools in hospitals; survey samples are self-selected; the trials measure different things (time vs output) and are not comparable one to one

The demo: AI reads the documentation, writes and tests an adapter, a person releases it, and then it runs with the model switched off

Live demo, explained in advance · Five steps, two systems · Demo design: own; systems shown are stand-ins

BEFORE THE DEMO

0 Set-up

Two systems that do not talk: a source system with a proprietary export and a thin vendor manual, and a target that accepts a standard format (e.g. HL7 FHIR)

WATCH FOR
The starting point: no interface, only documents

DESIGN TIME · AI ON the model does most of the work, a person releases it

1 AI reads the documentation

The coding agent reads the manual and sample messages and proposes a field mapping; it lists what the manual leaves open

WATCH FOR
Where it asks instead of guessing

2 AI writes adapter and tests

It writes the adapter and test cases on synthetic patients, runs them and fixes failures until the tests pass

WATCH FOR
Tests on synthetic patients turn from red to green

3 A person reviews and releases

Read the mapping and the test report, correct what is wrong, approve the release; the adapter is frozen as a version

HUMAN GATE

WATCH FOR
The release gate: a person, not the model

RUN TIME · AI OFF

4 The adapter runs, the model is off

The model is disconnected; new messages flow from source to target through the released adapter

WATCH FOR
Same input, same output, every time; no model in the path

DATA ON STAGE

Building and testing: synthetic patients only (generated, e.g. with Synthesia)

Running: new demo records the model has never seen; no patient data on stage

Model: a frontier model in the cloud is fine here, because no patient data reaches it

WHAT THE DEMO DOES NOT SHOW

A production interface: that needs security review, change control and monitoring

The interface vendor's licence terms, which can be the real bottleneck

Speed in a real hospital: no published measurement of an AI-built hospital interface found

RELEASE RECORD

Each release keeps the **tests that ran, the person who approved it and the version**, so the adapter can be audited and rolled back without the model

SOURCES ① Uebernickel, Build with AI, run without it: concept (own framing, 2026) · ② HL7 International, FHIR R4 specification (target format)

Further sources: Walonoski et al., Synthesia synthetic patient generator, JAMIA 2018 · Note: The demo runs on stand-in systems and synthetic or demo data; it illustrates the build-and-release pattern, it is not a measurement

The Verification Matrix sorts tasks by how complex they are and how hard they are to check, and gives each field its own form of human oversight

Concept · Four fields, two axes, five criteria each

VERIFICATION EFFORT: how hard is it to check? →

C

Proof Trail

OVERSIGHT AT MILESTONES · LOW COMPLEXITY · HIGH VERIFICATION

Claims stay assumptions

- The output looks simple, e.g. a risk score or an imaging flag
- Whether it is right shows only later, against outcomes
- People check at milestones, not every single output

A

Autopilot

OVERSIGHT AFTER THE FACT · LOW COMPLEXITY · LOW VERIFICATION

Act and undo

- AI acts on its own, e.g. booking an appointment slot
- People watch the log and correct mistakes

D

Co-Creation

CONTINUOUS OVERSIGHT · HIGH COMPLEXITY · HIGH VERIFICATION

Build together, release by gate

- Hard to produce and hard to check
- People and AI work on the output together
- Several roles are involved in confirming it
- Nothing leaves without explicit approval
- e.g. a treatment proposal prepared for a tumour board

B

Sign-Off

OVERSIGHT PER OUTPUT · HIGH COMPLEXITY · LOW VERIFICATION

Draft and approve

- AI drafts at volume
- Each item is quick to check against its source
- An expert approves every output, e.g. a visit note signed by the physician

TASK COMPLEXITY: how much expertise and context it takes →

How a task is placed

Five yes/no criteria per axis: 0–2 = low, 3–5 = high. Risk enters through the cost of correction

TASK COMPLEXITY

- 1 Needs licensed expertise
- 2 Combines several sources
- 3 Needs patient-specific context
- 4 Open-ended output (text, judgement)
- 5 More than ~10 minutes of skilled time

VERIFICATION EFFORT

- 1 Cannot be checked against a reference within minutes
- 2 Correctness shows only later
- 3 Needs more than one role to confirm
- 4 An undetected error reaches the patient and cannot be undone
- 5 Regulated device function or legal liability

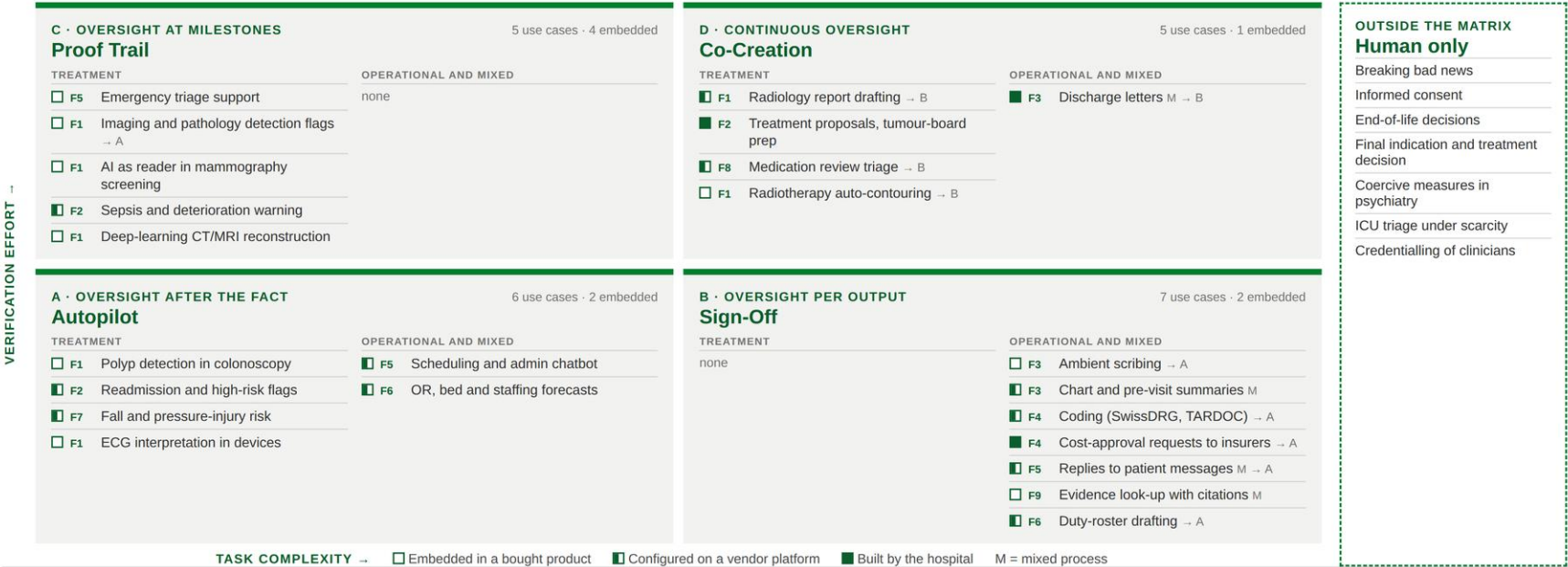
Human only

Outside the matrix by choice, not by technical limit: e.g. consent, delivering bad news, the final treatment decision

Drift: when checking gets cheaper (source links, structured output, rule checks, monitoring), tasks move down and left, towards Autopilot

About half of treatment AI comes embedded in bought products, mostly in Proof Trail; operational AI is mostly configured on vendor platforms

23 hospital AI use cases · Process type and design control: own assessment · → drift (hypothesis)



OUTSIDE THE MATRIX Human only

- Breaking bad news
- Informed consent
- End-of-life decisions
- Final indication and treatment decision
- Coercive measures in psychiatry
- ICU triage under scarcity
- Credentialling of clinicians

SOURCES ① Ueberrickel, F. (2026), Every task can be AI-supported: the question is how humans stay in charge, working note, University of St.Gallen · ② ASTP/ONC (HHS), Data Brief 80 (2025): US hospitals' predictive models by source, 2024: EHR developer 80%, third party 52%, self-developed 50%

Note: Design control classifies a deployment, not a use case: a sepsis model can be bought, configured in the EHR or built in-house

© 2026 Falk Ueberrickel