ACOFP-UMW · CLINICAL ESSENTIALS

AI for Clinicians

Andrew C. Bland, MD, FACP, FAAP

Save my contact

Scan to add my contact card

Download the slides

Scan to download all 66 slides (PDF)

urinenephrology.org/talks/

ACOFP-UMW · CLINICAL ESSENTIALS

AI for Clinicians

Practical AI for modern clinical practice

Andrew C. Bland, MD, FACP, FAAP
Medical Associates Nephrology

Validate · Understand · Constrain

Savery Hotel · Des Moines, IA · September 20, 2026

WHY THIS MATTERS

Who am I?

THE CLINICIAN

Board-certified in internal medicine, pediatrics, and nephrology. Chair of Nephrology at Medical Associates. MBA and MS in data science and predictive analytics. Practicing in Dubuque, Iowa. Adjunct Faculty, University of Dubuque PA Program. Associate Professor Internal Medicine and Pediatrics University of Illinois COM

THE BUILDER

Built a working AI operating system from scratch to help me see patients and educate students of Nephrology. Designed to bridge the implantation gap, keep current on literature, transform data into actionable information while making patient centered decisions without duplicative work.

The Dubuque Ice Harbor

Natural ice:

- harvested in winter
- stored in insulated icehouses
- cut into blocks
- and delivered to household iceboxes
- by wagon
- By 1900, ice harvesting was Maine's second-largest industry, shipping more than one million tons annually to markets as far away as India.
- In Iowa,1921 Cedar Falls Ice House alone could store 16 million pounds of Cedar River ice

Electric refrigeration shifted cooling from delivered ice to continuous, on-site refrigeration. By the 1950s, household block-ice delivery had largely disappeared, although commercial ice production continued.

Peak and Decline

- **In 1919, 63 Iowa establishments produced 324,844 tons—nearly 650 million pounds—of manufactured ice, valued at $1.9 million.** This excludes natural ice harvested from rivers and lakes.
- By the 1930s, widespread electric refrigeration rapidly dismantled Iowa's ice-harvesting and home-delivery economy.

WHY THIS MATTERS

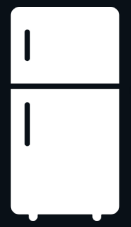
Don’t build a better ice box.

Same goal, two roads. One optimizes the old paradigm forever. The other changes the substrate.

THE ICE INDUSTRY — optimize the old paradigm

Harvest ice faster → **Thicker** insulation → **Better** delivery cart → **Machine-made** ice plants → **Better** ice box - - - still just **cold storage**

In medicine today: a better scribe · a better inbox · better data capture — a better icebox.



ELECTRICITY + ENGINEERING CONTROLS — change the substrate

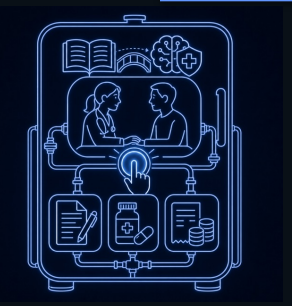
Same goal: Different Process
a different substrate — not a better icebox

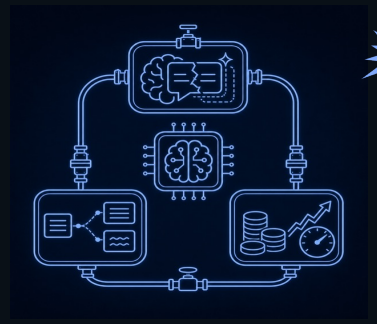
A new system
fresh food reliable at home without the mess of melting water or being home delivery

THE GOAL

Fresh · safe · edible

in medicine:
timely · safe · effective care





Medicine’s shift: reorient to the doctor–patient encounter · turn data into **actionable information.**

Optimization can perfect the old paradigm while missing the shift.

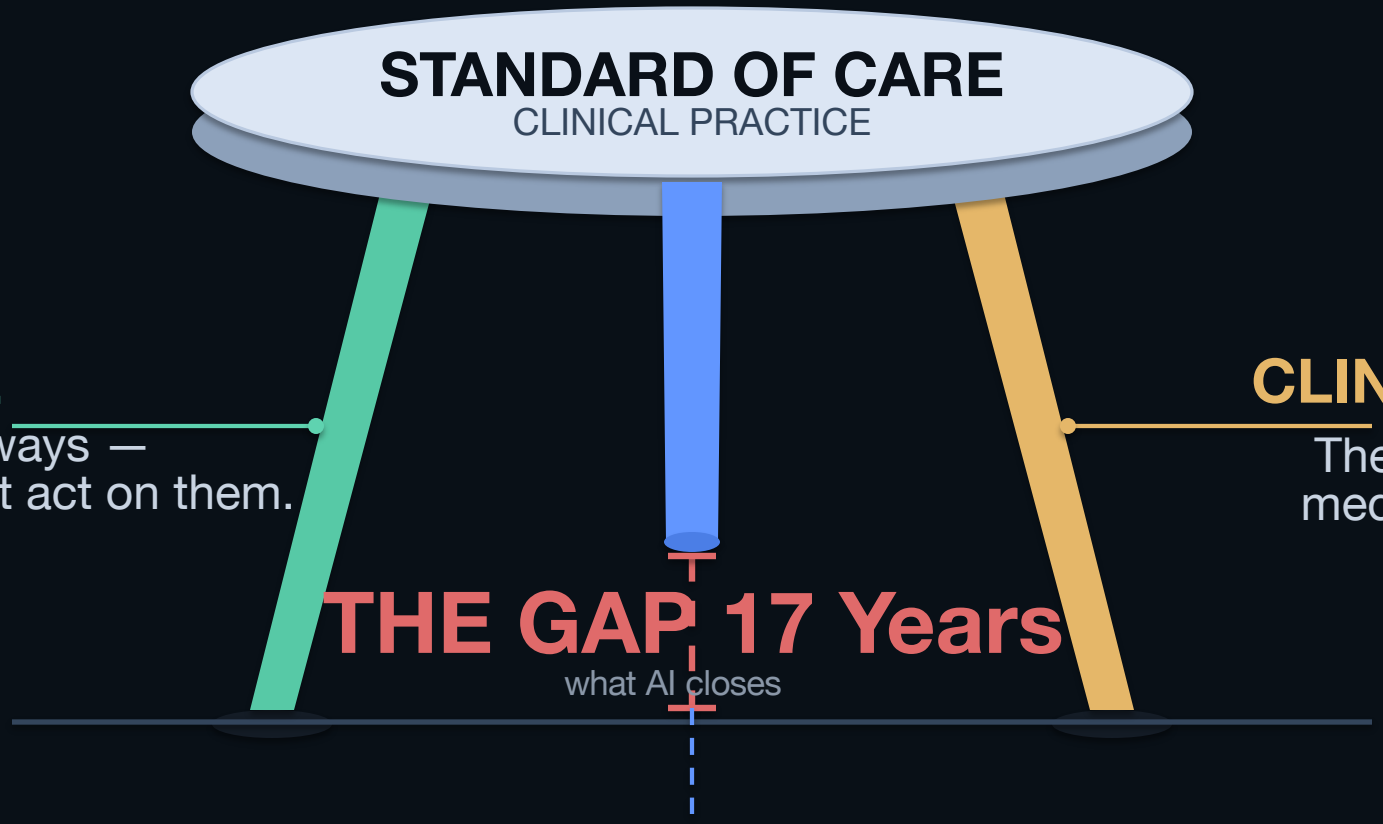
Benchmark against the goal — not the old workflow.

The Past

TRANSLATIONAL SCIENCE · THE IMPLEMENTATION GAP

Every standard of care rests on three legs.

Two are strong. The third — implementation — is where proven medicine stops short



BASIC SCIENCE

Discovery of the pathways — and the medicines that act on them.

CLINICAL SCIENCE

The evidence that those medicines actually work.

IMPLEMENTATION SCIENCE

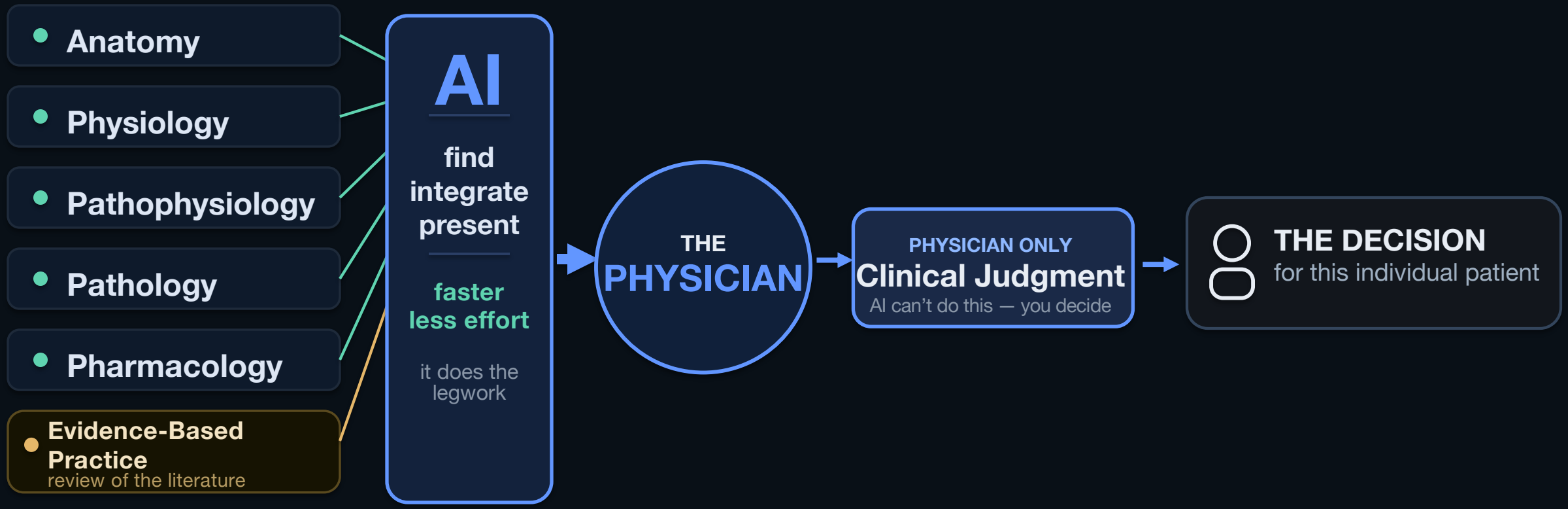
Getting proven care to **every** patient, reliably — the leg that comes up short. The seat can’t sit level until it’s built.

THE PHYSICIAN'S JOB · AI AS FORCE MULTIPLIER

AI gathers the evidence. The physician decides.

AI cuts the time and effort to find, integrate, and present the information — clinical judgment and the decision stay with the physician.

THE INFORMATION — EVERYTHING BUT JUDGMENT



AI is a force multiplier — it speeds finding, integrating, and presenting the evidence.
It does not decide. The physician does.

WHY THIS MATTERS

THE STAKES

Innovation improved ice delivery — not freshness

EHRs	Quality measures	Risk scores	AI scribes
Ice Factory	Better Ice Box	Thicker Insulation	Better ICE delivery

17 years

is a widely cited estimate for research findings to reach routine practice. AI may compress evidence finding and synthesis — not the entire implementation pathway.

Morris ZS, Wooding S, Grant J. J R Soc Med. 2011;104(12):510–520 · PMID 22179294

ENOUGH AI TO USE IT

The journey of medical knowledge

AI Cuts the time and energy needed to find the evidence and synthesize it

THE VOCABULARY · WHAT “AI” ACTUALLY CONTAINS

“AI” is three nested ideas — not one.

Machine learning lives inside artificial intelligence; deep learning lives inside machine learning. Modern generative AI usually means the innermost ring.

Artificial Intelligence

Any machine doing a task that would need human intelligence.

e.g. a chess engine, a routing app, Siri

Machine Learning

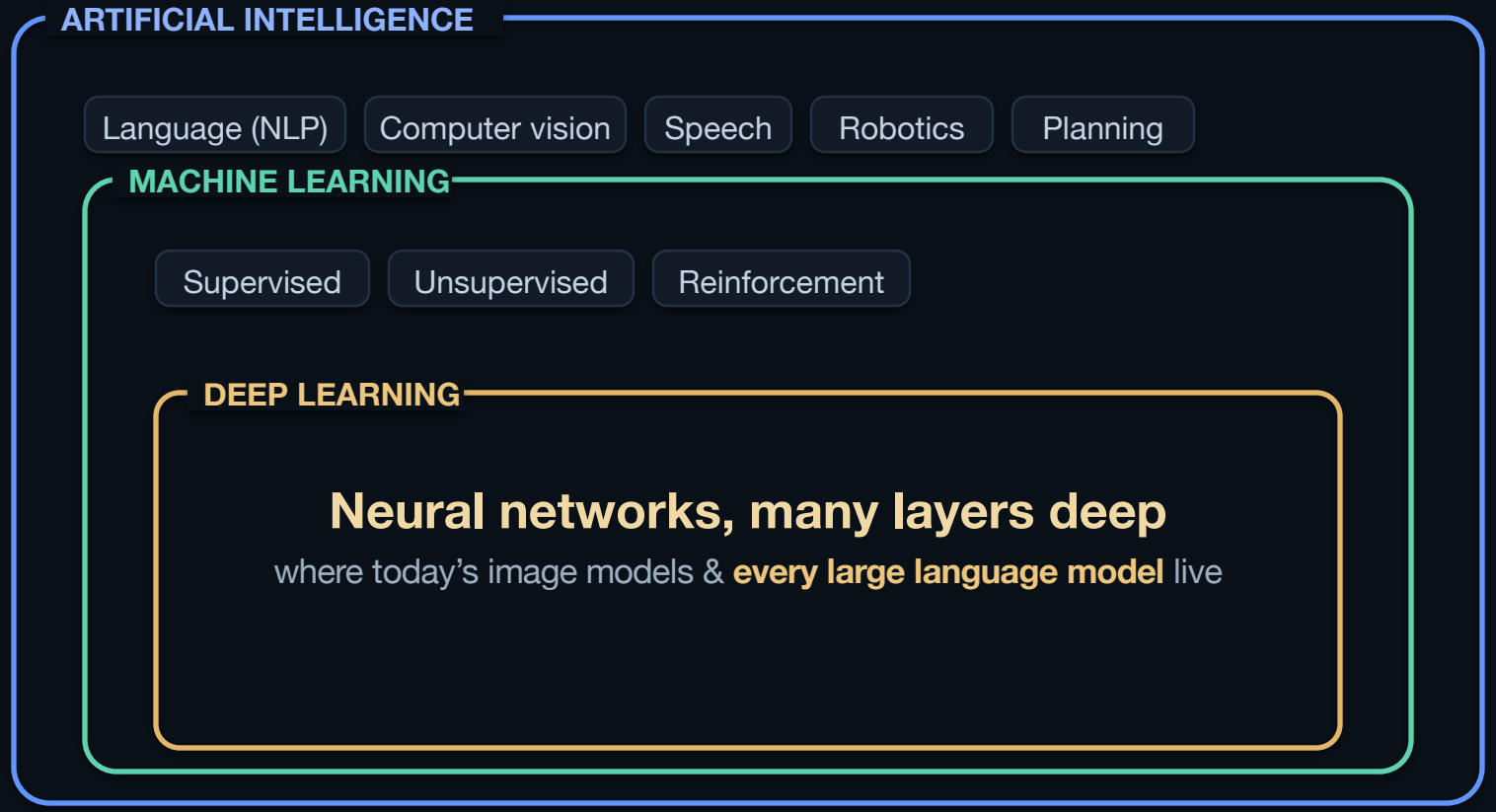
Learns the pattern from data instead of being hand-coded with rules.

e.g. a readmission risk score, a sepsis alert

Deep Learning

Many-layered neural networks that learn their own features from raw input.

e.g. imaging, ECG analysis, modern LLMs



WHY IT MATTERS For modern generative tools, “AI” usually means the innermost ring — deep learning.

THE VOCABULARY · WHERE CLAUDE, GPT & GEMINI COME FROM

Modern LLMs are deep learning — built on decades of AI.

The chatbots are the newest layer of a field that began with hand-coded logic. Each box below is a **subset of the one to its left** — and a leap forward in time.

① SYMBOLIC AI
1950s–1980s

Rules & logic

Coded by hand, step by step.

- LISP · Prolog
- Expert systems
- If-this-then-that

Brittle — knew only what we said.

→

② MACHINE LEARNING
1990s–2010s

Learns from data

Finds the pattern, not the rule.

- Regression · trees
- Support vector machines
- Risk scores

Improved as you fed it data.

→

③ DEEP LEARNING
2012 →

Neural networks

Features learned from raw input.

- Reads images, ECGs
- CNNs · transformers
- The 2012 breakthrough

Suddenly it could see & hear.

→

④ FOUNDATION MODELS
2020 →

One general model

Trained once; adapts widely.

- **LLMs:** Claude·GPT·Gemini
- **Generative:** images, video
- Deepfakes live here too

Reaching your exam room now.

THE THROUGH-LINE Today's generative AI sits in the far-right box — built on all the rest.

An LLM, in clinical terms

Language model	A learned system that predicts and generates token sequences
Large	Many learned numerical parameters and substantial training
Multimodal	Some models also accept images, audio, or video
General-purpose model	Can perform many tasks; it is not automatically a clinical device

OpenAI key concepts · Gemini model catalog

INSIDE AN LLM · FOUR DIFFERENT LEVERS

Tokens enter. Parameters work. Quantization makes them fit.

Tokens are temporary input. Vectors carry context. Parameters are learned machinery. Quantization changes precision — not the number of dimensions or parameters.

① TOKENS

TEMPORARY INPUT

Pieces of the prompt

Set context length and cost.

- Prompt + response
- KV-cache memory
- Gone after the request

More tokens ≠ more parameters.

② VECTORS

CONTEXTUAL STATE

Thousands of numbers

Meaning is distributed.

- Token → numeric vector
- Context reshapes it
- Ready for matrix math

No single axis means "city."

③ PARAMETERS

LEARNED MACHINERY

Learned numeric weights

Transform the vectors.

- Attention matrices
- Feed-forward weights
- Usually fixed in use

Training changes them.

④ QUANTIZATION

SAME MODEL, FEWER BITS

Coarser numeric storage

Cuts memory; may trim quality.

- **16-bit → 8-bit → 4-bit**
- **Same matrix dimensions**
- Runtime memory still extra

Fewer bits ≠ fewer parameters.

INSIDE AN LLM · TWO DIFFERENT PHASES

Training changes the weights. Your prompt does not.

Training predicts, measures error, and nudges billions of parameters.
Inference applies those learned matrices to new tokens — one forward pass at a time.

One architecture
two very different phases

Training
Changes the weights.

Predict
Guess the next token.

REPEAT
Compare with the answer

Adjust
Backpropagate the error.

OPTIMIZER
Nudge billions of weights

Learned weights
Store compressed patterns.

Vast token batches **repeat** the cycle.
Small corrections accumulate.

The final matrices encode **patterns** **associations**, styles, and relationships.
They are compressed knowledge —
not a searchable database.

Inference
Uses the weights.

Encode
Prompt tokens become vectors.

CONTEXT
Attention mixes relevance

Generate
Compute next-token probabilities.

REPEAT
Append one token and continue

THE DIVIDE Training updates parameters. Inference updates the context — not the core model.

UNDERSTAND · WHY HALLUCINATIONS HAPPEN

Fluency is not a truth test.

Four mechanisms explain how a model can sound right while being wrong.
The controls should target the mechanism — not the tone.

① NEXT-TOKEN OBJECTIVE
PRETRAINING

Predicts continuation

No true/false label for every claim.

- Most plausible token
- Not a database lookup
- Post-training can help

Optimizes continuation directly.

→

② WEAK EVIDENCE
PARAMETRIC MEMORY

Facts are distributed

Rare or current facts are harder.

- Patterns, not rows
- Limited provenance
- Retrieval adds evidence

Plausible ≠ retrieved or current.

→

③ FORCED SPECIFICITY
PRESSURE TO ANSWER

Forced specificity

Exact details can invite guessing.

- Exact dose
- Exact date
- Exact citation

Permit "insufficient evidence."

→

④ COMPOUNDING
AUTOREGRESSIVE

Errors recruit errors

Each token becomes new context.

- **Early mistake**
- **Coherent story**
- Invented support

Verify before clinical action.

THE CONTROLS Retrieve · allow abstention · require evidence · verify before clinical action.

INSIDE AN LLM · HIDDEN VECTOR SIZE

Local is a location — not a model size.

Qwen3 0.6B	=	1,024 dimensions
Llama 3.1 8B	=	4,096 dimensions
Llama 3.1 70B	=	8,192 dimensions
Llama 3.1 405B	=	16,384 dimensions

Hidden size = vector width. Sources: Qwen3 config; Meta Llama 3 paper. Proprietary frontier dimensions are not published.

Every bit costs memory—and can protect accuracy.

1-bit	8.82 GB — research frontier; not a generic drop-in Llama format	
3-bit	26.46 GB — aggressive compression; quality is method-dependent	
4-bit	**35.28 GB — capacity sweet spot; strong local balance when validated**	
8-bit	70.55 GB — quality sweet spot; often near baseline when tuned	QUALITY FIRST
16-bit	141.11 GB — reference fidelity; no quantization error	

Raw weights only: 70,553,706,496 × bits ÷ 8. Runtime adds scales, metadata, KV cache, activations, and work buffers.

The weight-memory calculation

Approximate model size	16-bit	8-bit	4-bit
8B parameters	16 GB	8 GB	4 GB
32B parameters	64 GB	32 GB	16 GB
70B parameters	140 GB	70 GB	35 GB
120B parameters	240 GB	120 GB	60 GB

Weights only · decimal GB · calculated estimates, not machine recommendations

CPU, GPU, and accelerators

Component	Main contribution	Example in this workflow
CPU	General logic and application work	Validate units; query a database
GPU	Many numerical operations in parallel	Run model inference
TPU / NPU	Specialized acceleration	Supported model operations
Memory	Holds the active data and weights	Keep the model and its cache resident

Apple Mac Studio specifications · NVIDIA RTX PRO 6000 Blackwell Server Edition

Current hardware examples

Configuration	Memory capacity	Advertised bandwidth
Mac Studio M5 Max 40-core GPU option	Up to 128 GB unified	614 GB/s
Mac Studio M5 Ultra 80-core GPU option	Up to 512 GB unified	1.2 TB/s
RTX PRO 6000 Blackwell Server Edition, one GPU	96 GB dedicated GDDR7	1,597 GB/s

Specifications checked 13 Sep 2026 · capacity ≠ speed ≠ clinical accuracy

Choose the smallest trusted system that can do the job.

DETERMINISTIC GATES

VERIFY

Calculations, citation checks, redaction, and permission rules. Auditable and repeatable.

LOCAL MODELS

PRIVATE + REPEATABLE

Sensitive, high-volume, stable tasks on controlled hardware — with real security controls.

FRONTIER MODELS

HARD REASONING

Complex synthesis or adjudication in a BAA-covered, approved data boundary.

Route by task, evidence burden, data sensitivity, latency, cost, and consequence. The strongest model is not automatically the safest workflow.

Current model families

Provider	Models to recognize	Role described by provider
OpenAI	GPT-6 Astra; GPT-5.6 Sol / Terra / Luna	Complex work; cost / speed choices
Anthropic	Fable 5.1; Opus 5; Sonnet 5; Haiku 4.5	Reasoning and agentic work; faster tiers
Google	Gemini 3.8 Flash; 3.5 Flash-Lite; 3.1 Pro preview	Agents and coding; throughput; preview
Open weights	Qwen3.6-35B-A3B	Deployable multimodal MoE example

Dated snapshot · provider descriptions are not independent clinical validation

ENOUGH AI TO USE IT

The AI maturity framework

L6 AI Agents and Programming

L5 Multi-model orchestration- Builds off of Truth

L4 Knowledge Hub- Generates Local validated truth

L3 **Co-worker — Doximity Ask, Update AI, Claude or Chat GPT Agents**

L2 Chat / assistant — UpToDate, Open Evidence

L1 Prompting — one ask, one answer

Your system should outlive a model release

Keep the clinical rules, evidence, and tests outside the model

Treat a new model as a candidate replacement: run the same cases, compare failures, and promote it only when it meets your requirements.

Anthropic · Clinical pipeline architecture, inspected 2026-09-13

LOCAL REDUCES DISCLOSURE — NOT DUTY

HIPAA follows the PHI path—not the logo.

No-PHI / uncovered account

Standard Claude accounts are not BAA-covered.

Free · Pro · Max · Team · Cowork ≠ covered

Approved deployment

OpenEvidence: BAA in Terms.
Anthropic: HIPAA-ready Enterprise/API + signed BAA.

Verify feature, account, workflow, and policy

Local / on-premises

Removes one cloud disclosure, but not the Security Rule.

Encrypt · restrict · log · patch · back up · assess risk

HHS: a cloud vendor handling ePHI on your behalf requires a compliant BAA plus your own safeguards. Vendor terms checked 7/12/2026 — recheck before use.

EVIDENCE WITHOUT SURRENDER

TWO 2026 STUDIES

Related questions. Different study designs.

These evaluations were related — but not a direct replication.

	Nature Medicine · peer-reviewed	OpenEvidence · preprint
Queries	100 real clinician queries	620 platform-derived queries
Reviewers	12 · one institution	149 · specialty-matched · 36 states
Method	task-specific rubric	blinded pairwise preference
Finding	**frontier models led**	**OE net win margin: +25–39 pts**

Query source, comparators, model versions, and scoring design materially shape rankings.

EVIDENCE WITHOUT SURRENDER

Read the provenance — not just the winner.

NATURE MEDICINE

Peer-reviewed. 100 real clinician queries. Frontier and specialized tools evaluated with task-specific rubrics.

OPENEVIDENCE PREPRINT

620 platform-derived queries. Specialty-matched, blinded pairwise ratings. Prespecified analysis.

IMPORTANT CONTEXT

OpenEvidence co-developed the plan, ran recruitment and data collection, and paid respondents. Context informs interpretation; it does not invalidate a result by itself.

The catch — a real case

Hyperosmolar hyponatremia from acute ethanol

ODS risk during correction

Does fludrocortisone add meaningful psychiatric risk?

I caught the problem because I had this literature in my head. In most scenarios, most of us won't. That's the point of what comes next.

EVIDENCE WITHOUT SURRENDER

The danger is confidence without completeness

Open Evidence — in its own words

“I gave you confident recommendations with citations as if the evidence directly applied. It didn't. I dressed up extrapolation as established fact.”

Screenshot slot: Open Evidence

The frame isn't “it lies” — the blinded study found hallucination parity. The danger is a confident but **incomplete** clinical picture.

EVIDENCE WITHOUT SURRENDER

Real citations can still create a misleading picture

Open Evidence — in its own words

“FDA label warnings are class-based boilerplate… the psychiatric data comes almost entirely from glucocorticoids at immunosuppressive doses.”

Screenshot slot: Open Evidence

It didn't fabricate — it selectively assembled **real citations into a misleading picture.** More dangerous than fabrication: this only catches you if you already know.

EVIDENCE WITHOUT SURRENDER

Why a model can sound more certain than it is

A MODEL SELF-CRITIQUE — A CLUE, NOT EVIDENCE

“I default to a confident recommendation… Citations can make an answer feel more complete than it is… I may understate uncertainty and treat extrapolation as established fact.”

Screenshot slot: model self-critique

A confident answer can omit uncertainty. Build the workflow that catches it.

START WHERE RISK IS LOW

What AI is / isn't ready for

READY

- Drafting
- Summarizing
- Structuring
- Pattern recognition

NOT READY

- Autonomous clinical decisions
- Replacing the encounter
- Real-time EMR data
- Anything unreviewed

START WHERE RISK IS LOW

Which tool for which task

TASK	RECOMMENDED TOOL	DATA BOUNDARY
Patient education	Doximity Ask / Claude	Approved account · or de-identify
Prior-auth appeal	Doximity Ask	Approved account
Reference / fact check	OpenEvidence → UpToDate → PubMed	No PHI
Documentation	Approved ambient scribe / EHR AI	Approved account
Complex synthesis	Frontier model + verification stack	De-identify · or approved

START WHERE RISK IS LOW

Doximity Ask — a low-friction start

APPROVED TOOL, LOW-RISK TASK

- Confirm your organization’s approved deployment
- Use the exact BAA-covered account when PHI is involved
- Draft prior auths, referrals, or patient education

A useful draft in 10 minutes

Open the approved tool, draft one letter, then review every line.

PROMPTING → AGENTS → SOFTWARE

The “death of prompting”

The clever prompt stops being the whole product

Clear instructions remain. The durable advantage moves toward context, tools, feedback, and a system that can finish the job.
Astra 6, Opus 5, Fable 5.1 have changed the rules on prompting

A useful prompt is a compact specification

Goal	What outcome do I need, and for whom?
Inputs	Which sources may it use? What is missing?
Boundaries	What may it do, and where must it stop?
Evidence of success	What should the output contain, and how will I check it?

An agent works best in a loop

Observe → choose → act → inspect the result

The model decides the next permitted step, uses a tool, and adjusts based on what happened. A stopping rule keeps the task bounded.

Delegation needs a return contract

Assignment	A bounded task with explicit permitted inputs
Result	Structured observations or an artifact
Provenance	Sources, dates, and what changed
Uncertainty	Missing, unreadable, or conflicting evidence
Completion	What passed; what still needs a person

Computer use is one kind of tool

An agent can operate a screen when a useful API is absent

It must confirm the screen, take a bounded action, and verify the new state. A click alone is not proof that the task succeeded.

Software development starts with a clinical specification- The logic isn’t dependent on AI

Describe the friction	What forces the clinician to reconstruct the story?
Show an example	A synthetic case and the output you want
Specify behavior	Correct output, missing-input behavior, and stopping points
Build a small slice	One pathway with a visible beginning and end
Review it in use	Check the behavior with clinicians before expanding

AI-generated code needs ordinary engineering

Check	What it should establish
Unit and data tests	Equations, units, dates, and missing-data behavior
Workflow tests	The right task reaches the right person
Adversarial cases	Bad inputs fail visibly and safely
Versioning and rollback	A change can be traced and reversed
Clinical review	The behavior fits the intended use

PROMPTING → AGENTS → SOFTWARE

The result can be a small, durable application

A program is a reusable clinical decision aid

One reviewed calculation, one consistent handout, one visible follow-up queue—available without rewriting the prompt each time.

START WHERE RISK IS LOW

Real workflow examples- You can do these as 1 offs Level 1-3 or build something more durable- Level 4

TASK	BEFORE	AFTER
Patient-education letter	20 min	**2 min**
Prior-auth appeal	45 min	**5 min**
Complex chart summary	15 min	**3 min**
Literature summary · differential brainstorm (then verify)	—	**minutes**
CME quiz · coding review	—	**minutes**

None of these is a clinical decision. Writing and structuring — where AI is strongest and the risk is lowest. Start here.

START WHERE RISK IS LOW

Prior auth — before / after

BEFORE

7 steps · 45+ min

The one on your desk you haven't started — because you don't have 45 minutes.

→

AFTER

6 steps · 5 min

Draft from the denial + your notes, then you review and sign.

100+ hrs

Illustrative assumption: 3/week × 40 min = 2 hr/week = 100+ physician-hours per year.

Live: a synthetic CKD patient-education letter

Physician review = step 1

Synthetic details only. I generate — then I review every line.

START WHERE RISK IS LOW

Level 4 The Knowledge Hub
The reference verification stack

Verified sources reduce risk. Physician review still closes the loop.

START WHERE RISK IS LOW

Make critique part of the workflow

A TOOL THAT ASSERTS

Act 1: a confident answer with the uncertainty buried.

→

A SYSTEM YOU INTERROGATE

“What does the evidence actually support? Where is this extrapolation? What are the methodological weaknesses?”

Graded against evidence · science · methodology.

Run the same fludrocortisone case that fooled it — now it must say what's solid vs. extrapolated. A system you can interrogate beats a tool that only asserts.

START WHERE RISK IS LOW

HIPAA AI zones

No-PHI / unapproved account

Do not enter PHI. Paid does not mean BAA-covered.

Use only formally de-identified information

Approved vendor account

PHI only in a BAA-covered product/account, under policy.

A BAA is necessary — not sufficient

Local / on-premises

Can reduce disclosure, but is not automatically HIPAA-compliant.

Verify telemetry, backups, egress, access, and logging

Check the exact product, account, BAA, configuration, permitted use, and organizational policy — not merely the brand name.

HYBRID PIPELINE · SYNTHETIC DATA

Scattered data → a control-charted pre-visit sheet.

PHI-bearing notes and identity remain local; de-identified inputs may enter an approved cloud-reading lane: deterministic code joins, validates, trends, and renders the pre-visit scorecard.

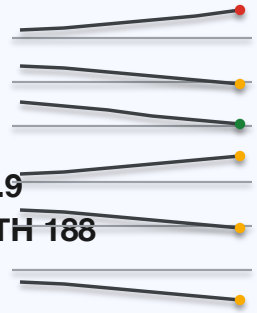
BEFORE · THE CHART
meds · labs · vitals, as they arrive

- Losartan 100
- K+ 5.4 (3wk fax)
- BP 148/86 · 142/84 · 138/80
- eGFR 38↓ · Cr 1.9
- Wt 88.5→86.4
- UACR 610→420
- A1c 7.8 (outside)
- Empagliflozin — when?
- Finerenone

Document extraction ▼ **Clinical synthesis** ▼ **Deterministic trend engine** →

10:40 CKD G3b · A3 PCP: Dr. Rivera **FLAGS: K↑ · HCO3↓**

▣ LABS — PRE-VISIT XmR TRENDS



AKI	Cr 1.9	⚠ HIGH (base 1.6)
CKD	eGFR 38	declining
CKD	UACR 420	✓ improv –31%
K / Mg	K 5.4 · Mg 1.9	⚠ HIGH
MBD	HCO3 21 · PTH 188	⚠ LOW
BP	138/80	target <130/80

GDMT Losartan 100 · Empagliflozin 10 · Finerenone 20 · Semaglutide 1/wk
+HTN / diuretic / binder Amlodipine 10 · Furosemide 40 · Patiromer 8.4 g
Lipid / MBD / anemia / GI Atorvastatin 40 · Calcitriol 0.25 · Ferrous sulfate 325 · Pantoprazole 40

▣ ASSESSMENT & PLAN

CKD G3b (A3), diabetic. Active signals: K↑ · HCO3↓; eGFR down, UACR improving.

▸ **Suggested: start alkali (HCO3 <22); recheck K before uptitrating finerenone.**

The journey through the clinical pipeline

Capture	Verified screenshots or an authorized structured source
Protect and read	Local privacy checks; bounded extraction
Validate	Values, units, dates, identity, completeness
Apply evidence	Versioned rules and explicit review states
Deliver and follow	Clinician view, patient information, and tracked work

Documented pipeline plus proposed follow-through layer

The local / cloud boundary

Work	Location in the documented design
Raw identity and clinical notes	Local
Privacy checks and identity join	Local
Cleared labs, meds, and vital grids	Permitted frontier-reading lane
Validation, clinical rules, rendering	Deterministic local code
Patient artifacts and structured corpus	Controlled private delivery

Clinical pipeline architecture, inspected 2026-09-13

A value needs more than a number

Field	Synthetic example
Observation	eGFR = 38
Unit	mL/min/1.73 m^2
Date	2026-09-01
Provenance	Cleared lab frame + extraction version
Review state	Readable; date and unit verified

Synthetic example · not a patient record

Validation is clinical infrastructure

Identity	Do all components belong to the same scheduled person?
Completeness	Were all expected sections accounted for?
Consistency	Do dates, units, and repeated values agree?
Plausibility	Does an implausible result trigger review?
Provenance	Can every output be traced to the same capture generation?

Clinical pipeline architecture, inspected 2026-09-13

The clinician’s pre-visit view

What changed	Longitudinal observations and medication changes
What is unresolved	Conflicts, absent inputs, and pending work
What may need action	Evidence-linked items for clinical review
Where it came from	A traceable path back to the observations

Clinical pipeline architecture, inspected 2026-09-13 · Scorecard v7

The patient’s information has a different job

Explain	What the reviewed findings mean in plain language
Agree	The decisions made together during the visit
Act	What to do, who will help, and when
Follow up	How the patient will know what happens next

Clinical pipeline architecture, inspected 2026-09-13

A report is an output. A loop needs state.

Order → result → review → action → confirmation

An application closes the operational loop only when someone owns the next step and the system can distinguish pending work from completed work.

Follow-through is an implementation goal; integration must be demonstrated

CARDIORENAL GDMT · THE SINGLE PATHWAY

One pathway — the scorecard, distilled.

The clinical scorecard reduced to the line that matters: risk → evidence → therapy → outcome.

① THE PATIENT AT RISK

Diabetic CKD

linked kidney and cardiovascular risk

eGFR 38 · stage G3b

UACR 420 · stage A3

K trending up · HCO3 low

High risk: kidney failure + CV death.

② THE EVIDENCE

The evidence defines it

each pillar, its own RCT

SGLT2i
DAPA-CKD · EMPA-KIDNEY · CREDENCE

ns-MRA
FIDELIO-DKD · FIGARO-DKD

GLP-1 RA
FLOW

RAS blockade
RENAAL · IDNT

Landmark, guideline-anchoring trials.

③ THE PATHWAY — GDMT

Evidence-based therapy

KDIGO-aligned · layered

1 ACE-i / ARB

2 SGLT2 inhibitor

3 Finerenone (ns-MRA)

4 GLP-1 RA (semaglutide)

Start early, layer up, monitor K.

④ THE OUTCOME

What it targets

for eligible patients

slower CKD progression

fewer CV / HF events

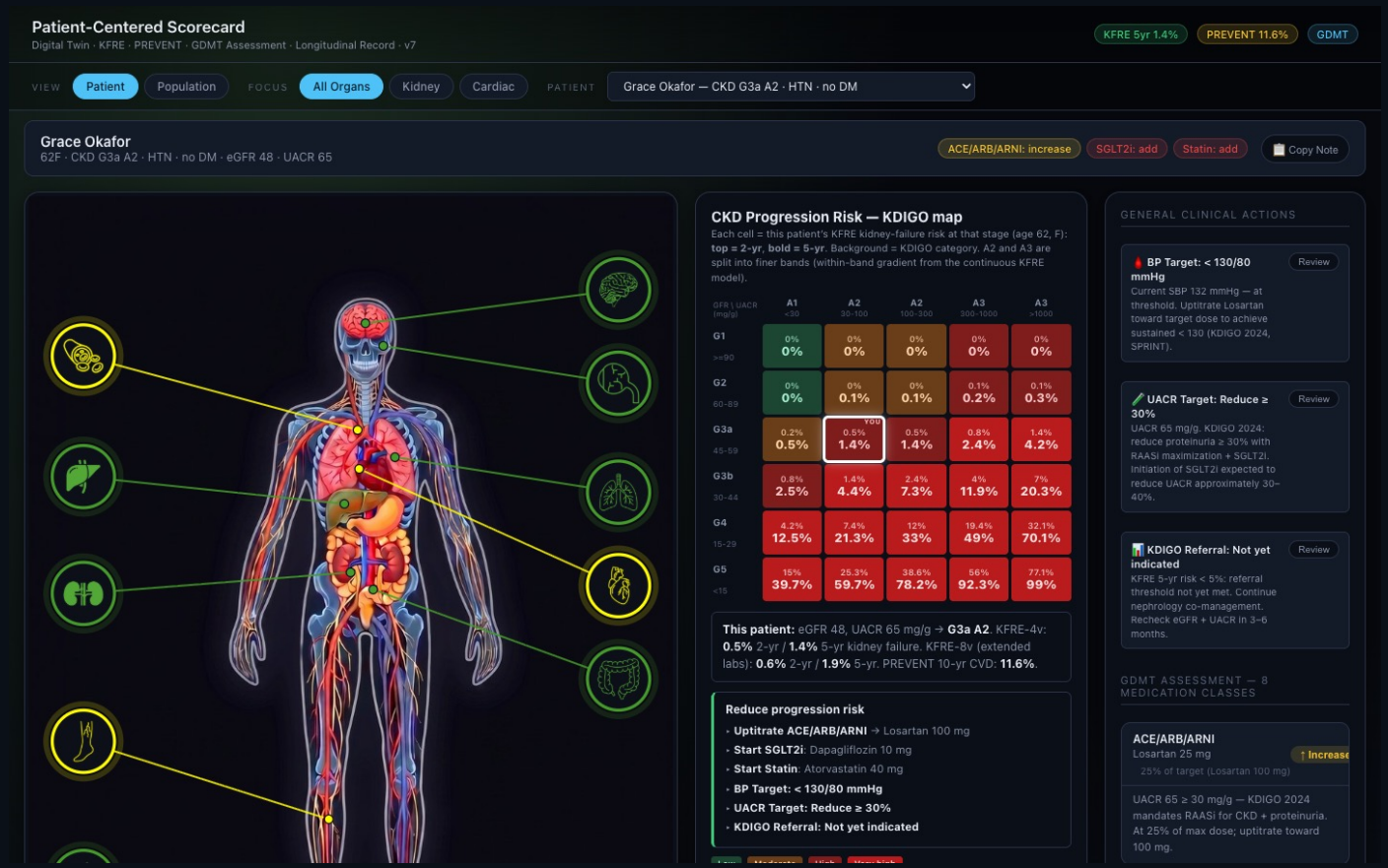
lower albuminuria

benefit varies by therapy

Use the therapies indicated for this patient.

THE IMPLEMENTATION LEG Class-specific trials — not one four-drug trial. Use indicated therapies; monitor kidney function and K^+.

The digital scorecard: one patient view

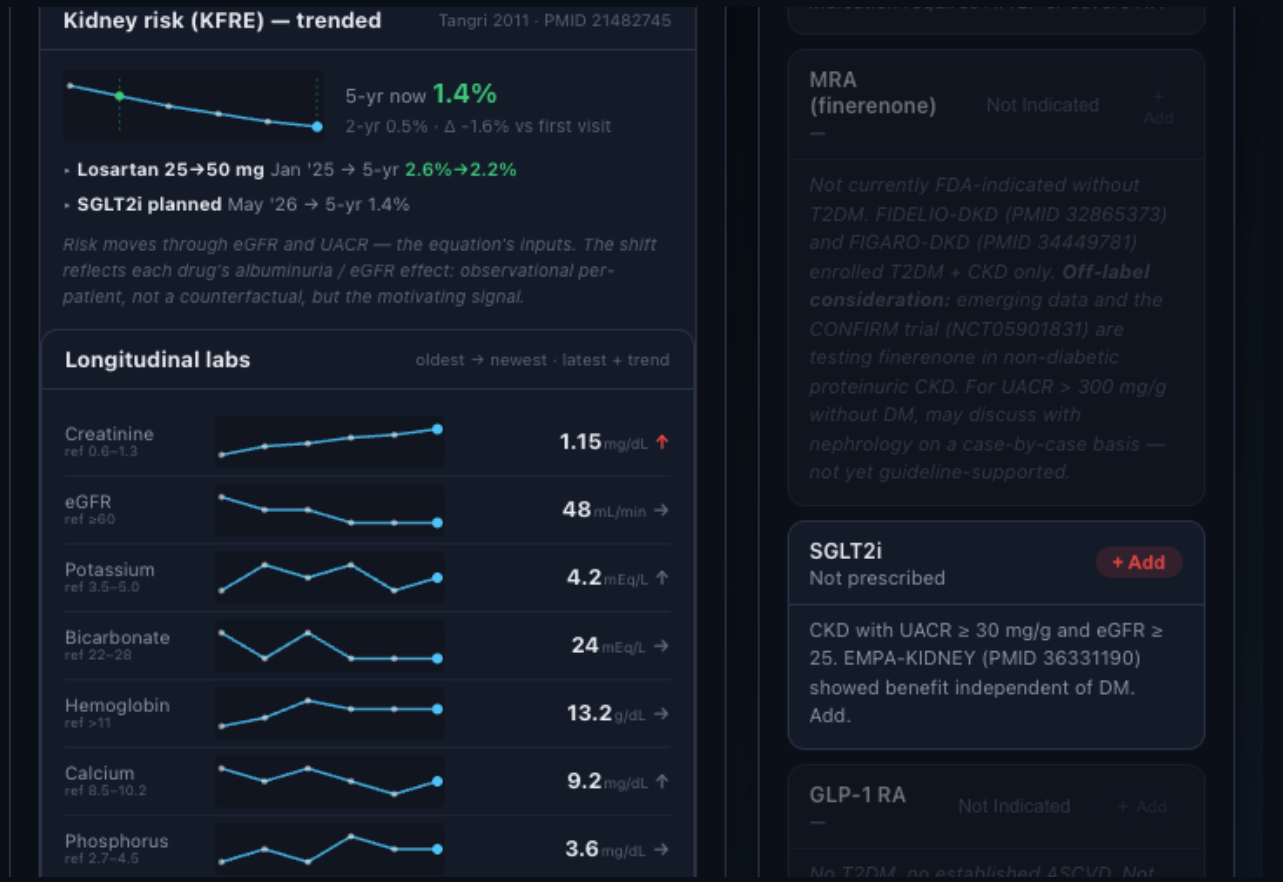


Risk framing, longitudinal information, and review items share a screen

Existing v7 prototype · synthetic patients · interface demonstration

Validate · Understand · Constrain

The patient dashboard: a longitudinal view



A record can show change over time and the work still pending

Existing v7 prototype · synthetic data · population feed not connected

Validate · Understand · Constrain

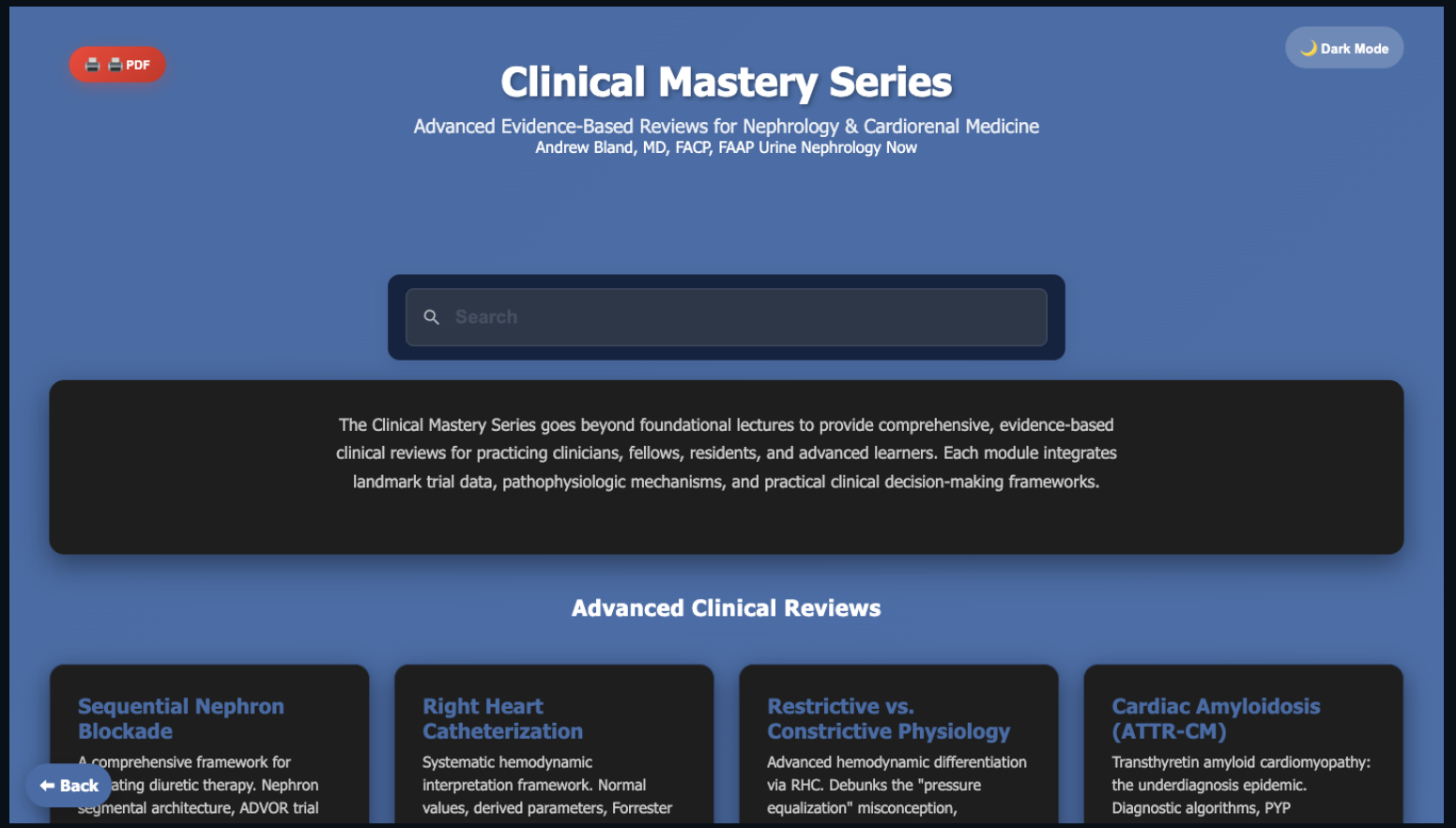
Urine Nephrology Now

A clinical question can become a reusable teaching resource

Live public website · urinenephrology.org

Validate · Understand · Constrain

Clinical Mastery: a shared evidence resource



The same reviewed knowledge can support teaching and clinical workflow

Public site captured 13 Sep 2026

Validate · Understand · Constrain

WHAT SAFE SYSTEMS CAN BECOME

Imagine: the integrated clinical companion

KNOWS THIS PATIENT

Every med, lab, symptom, and problem — structured. The longitudinal clinical view.

+

HAS READ THE LITERATURE

The current evidence for exactly those problems — living documents, kept current.

“Given her CKD and this new potassium, what changed this year — and what would you reconsider?”

The pieces exist today. The fully integrated companion is where this points. Any live version must stay inside an approved data boundary.

Clinical care should set the requirements

Stop making the clinical workflow fit the software

Start with the patient’s need. Make the data, interface, and automation support the clinical sequence.

The goal is care that follows through

Ask a question.
Build the system that helps act on the answer.

Evidence → patient context → clinical judgment → action → follow-up

Build your own

THE REINFORCING STRUCTURE

- CLAUDE.md or Agendt.md — rules / memory
- Prompts — the ask
- Artifacts — projects / templates / context
- Agents — delegation
- Hooks — guardrails + verification

rules → prompt → artifact → hook-verified → context

GO DEEPER

- Dex as AI Chief of Staff — one orchestrator
- OpenMed Agent — the market arrived at the same architecture
- Mac Sparky — “let the robot do the donkey work”
- Tom from ICOR — building AI swarms