Sanjeevani

Sanskrit: 'that which gives life'

Neuro-Symbolic AI for Automated CKD Triage & EMR Drafting via 1D NMR

Abhiram Radha Krishna · Kaggle MedGemma Impact Challenge · 2025

The Problem in Numbers

850M people with CKD	**90%** undiagnosed early
3hrs daily EMR burden	**0** LLMs can read NMR

Built entirely from scratch for this challenge.

Google's HAI-DEF models deployed strictly within their competence domain.

THE PROBLEM

A Silent Epidemic & An Interpretability Bottleneck

The Spectroscopic Bottleneck

^{1}H NMR spectroscopy of urine is the gold standard for early CKD metabolic profiling. It detects:

Citrate	depleted in renal tubular dysfunction
TMAO	elevated uremic toxin
Lactate	elevated in renal hypoxia
Creatinine	normalisation anchor metabolite

BUT: requires highly trained spectroscopists — a critical bottleneck for under-resourced clinics.

Why Pure LLMs Fail Here

LLMs are text-prediction engines. They cannot perform spectral deconvolution on raw float arrays.

- ✕ Hallucinate diagnoses from random noise
- ✕ Cannot extract metabolite ratios reliably
- ✕ No deterministic mathematical grounding
- ✕ Dangerous false confidence on bad data

Solution: Use LLMs only for what they're good at. Build a deterministic physics layer first.

ARCHITECTURE

Three AI Agents — Each Doing Only What It Can Do Well

STAGE 1

PaliGemma

paligemma-3b-mix-224

Visual Quality Control

Renders 2.5–4.5 ppm anchor region as image. Classifies 'sharp peaks' vs 'flat noise'. Halts pipeline on degraded samples.

STAGE 2

SanjeevaniEngine

Custom PyTorch CNN-Transformer

Biomarker Deconvolution

5-channel independent decoders extract creatinine, citrate, lactate, TMAO, taurine. Outputs deterministic risk score.

STAGE 3

MedGemma

medgemma-1.5-4b-it

Clinical Communication

Receives verified ratios only — never raw data. Dual persona: Physician Assessment & Plan + patient-friendly EMR draft.

🔒 Safety Gate: MedGemma never sees raw spectral data — only three verified decimal ratios from the physics engine. This prevents hallucinated clinical reasoning.

LIVE DEMO — SAFEGUARD

Patient C: Degraded Sample → Pipeline Halts Safely

Without Sanjeevani

1. Raw noise fed directly to MedGemma
2. LLM pattern-matches text from noise
3. Generates a plausible-sounding diagnosis
4. Clinician acts on a hallucinated result
5. Patient harmed — false negative discharge

With Sanjeevani

① PaliGemma Vision Gate

Renders anchor region (2.5–4.5 ppm). Classifies image as 'flat noise' — not 'sharp peaks'.

② Mathematical Creatinine Gate

Checks creatinine anchor amplitude < 0.01 threshold. Secondary biochemical validation — dual-layer QC.

③ Pipeline Halts

MedGemma is never invoked. Dashboard shows: 'INCONCLUSIVE (QC FAILURE) — Biomarker Math Halted: Data integrity compromised.'

[See dashboard screenshot: inconclusive_fig_1.png + inconclusive_fig_2.png]

DEMO — CLINICAL CO-PILOT

Patient B: Type 2 Diabetic → HIGH RISK DETECTED

Biomarker Panel

Creatinine **0.90** (anchor)

reference peak — normal

Citrate Ratio **0.06** / Cr

⚠ *severely depleted — renal tubular dysfunction*

TMAO Ratio **1.14** / Cr

⚠ *elevated — uremic toxin accumulation*

Lactate Ratio **1.36** / Cr

⚠ *elevated — anaerobic metabolism*

RISK SCORE: 3.22 Threshold: 1.2 → HIGH RISK

MedGemma — Physician View

"The depleted citrate ratio, elevated TMAO ratio, and elevated lactate ratio are indicative of metabolic acidosis and potential kidney dysfunction. These findings suggest impaired renal bicarbonate excretion... A renal ultrasound is recommended."

MedGemma — Patient EMR Draft

"Hi [Name], I'm following up on your recent check-up. Your fatigue, along with recent changes in your kidney function, is something we need to look into further... We're recommending a renal ultrasound."

TECHNICAL DETAILS

SanjeevaniEngine: Architecture & Performance Evolution

Architecture

Encoder	2× InceptionBlock (kernels 1,3,5,7) + MaxPool1D → 64-channel latent
Sequence	2-layer TransformerEncoder (d=64, heads=4) for long-range dependencies
Decoder	5 INDEPENDENT branches — one per biomarker, with BatchNorm + weighted loss
Loss weights	TMAO 5×, Lactate 3×, Citrate 2×, Taurine 2×, Creatinine 1×
Training	10,000 synthetic pseudo-Voigt spectra · 25 epochs · Adam optimizer

Performance Evolution

Version	Acc	Sens	Spec
V1 Baseline (shared decoder)	93%	96.6%	87.8%
V2 Indep. decoders + decoys	91%	96.1%	85.7%
Final: weighted loss (all 5 ch. active)	89.5%	100%	76.9%

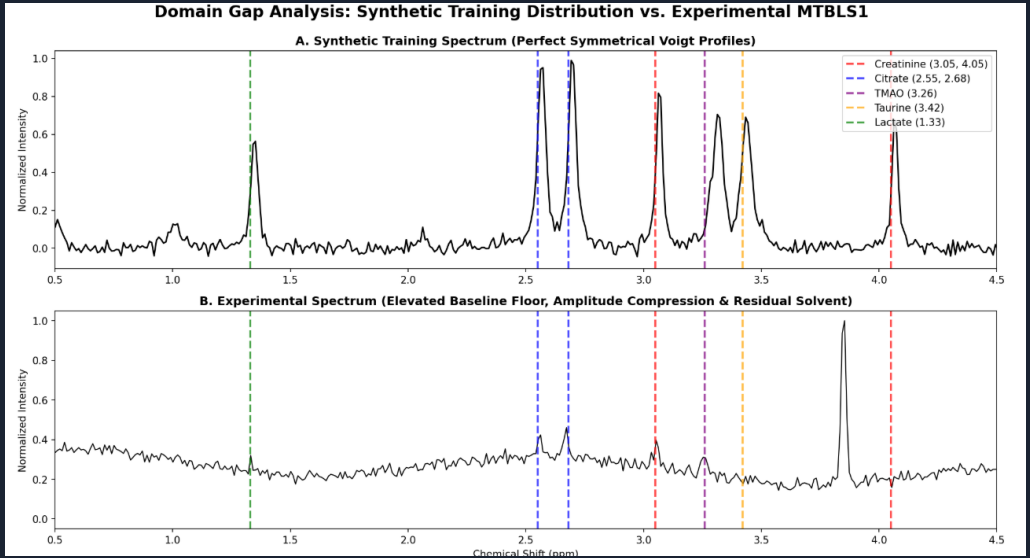
↓ *Accuracy intentional: reflects removing inflated closed-world metrics & forcing all 5 channels active*

[See: confusion_matrix.png — 100% sensitivity, 14 QC flags / 100 patients]

REAL-WORLD TRANSFER

MTBLS1 Zero-Shot Transfer: Honest Domain Gap Analysis

Domain Gap Analysis Figure



Two Domain Gaps Identified

Gap 1: Amplitude Mismatch

Synthetic: creatinine always dominant (0.8–1.2). Real urine: creatinine can be subdominant. Model outputs zero for creatinine channel.

Gap 2: Elevated Baseline Floor

MTBLS1 spectra show macromolecular background at 0.25–0.35 normalised intensity. Synthetic data has near-zero baseline. All real peaks compressed.

V2 Fixes: (1) Add broad Lorentzian baseline augmentation to generator · (2) Calibrate creatinine amplitude distribution from real MTBLS1 · (3) Spinach quantum spin Hamiltonian simulator for J-coupled multiplets

Sanjeevani

Neuro-Symbolic AI for Automated CKD Triage & EMR Drafting via 1D NMR

Dual-Layer QC

PaliGemma vision + creatinine biochemical gate prevent bad data reaching clinical reasoning

Channel Collapse Fix

Independent decoder branches + weighted loss forces all 5 biomarker channels to function

Hallucination Prevention

MedGemma receives only verified decimal ratios — never raw spectral data

Honest Transfer Eval

Zero-shot MTBLS1 test with quantified domain gap analysis and specific V2 roadmap

"A model that fails honestly and explicably is more valuable for clinical AI than one that succeeds silently on its own test set."

Abhiram Radha Krishna · Kaggle MedGemma Impact Challenge · github/kaggle notebook link