EMNLP 2026 · Findings · Poster

HiMed: Incentivizing Hindi Reasoning in Medical LLMs

Dingfeng Jiang* · Han Yan* · Chenze Ma* · Amit Kumar Jaiswal* · Ang Li · Yunxiang Jiang · Xinlei Xiong · Juhao Liang

Hongru Xiao · Xiang Li† · Fan Bu · Jiale Han · Ruchir Gupta · Prayag Tiwari† · Benyou Wang†

[1]CUHK-Shenzhen [2]IIT (BHU) Varanasi [3]Tongji University [4]Shenzhen Research Institute of Big Data [5]Shenzhen Loop Area Institute [6]Halmstad University

* Equal contribution † Correspondence: lixiang@sribd.cn · prayag.tiwari@ieee.org · wangbenyou@cuhk.edu.cn

Scan for code, data, and weights

github.com/FreedomIntelligence/HiMed

01 Overview

Native Hindi data and a decaying language-form reward improve medical reasoning without sacrificing English performance

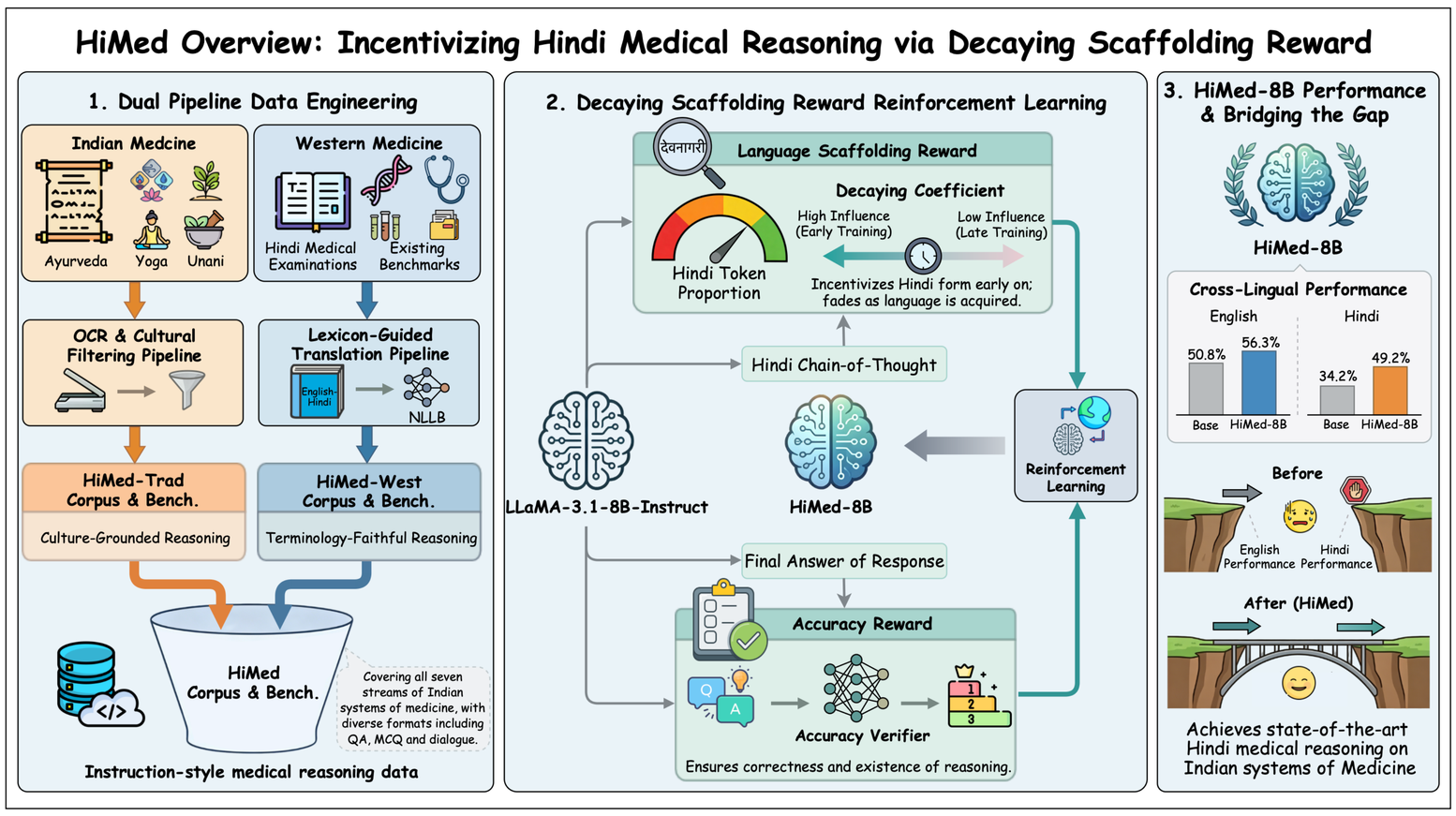


Figure 1. HiMed combines native and terminology-faithful Hindi data with decaying language scaffolding during reinforcement learning.

What changes with HiMed

- Hindi-native evidence. Two pipelines build reasoning data for Western medicine and all seven Indian systems of medicine.
- Form-aware reinforcement learning. DSR-RL rewards Hindi reasoning early, then decays that reward so accuracy governs the final policy.
- Cross-lingual gains. HiMed-8B improves Hindi accuracy while preserving English reasoning and domain terminology.

403K training instances — HiMed-Trad + HiMed-West

8.2K benchmark items — zero passage overlap

45,902 En-Hi terms — lexicon-guided translation

Headline result

HiMed-8B reaches 76.0% on HiMed-Trad, outperforming GPT-4o by 14.0 points.

The average English-Hindi gap falls from 16.6 to 7.1 points relative to the LLaMA-3.1-8B base model.

Expert-audited data quality

98.0% OCR acceptance — 200 audited segments

99.5% generation acceptance — 200 audited instances

97.0% translation acceptance — lexicon-guided NLLB

02 Main results

Zero-shot accuracy (%) across Western and Indian medical reasoning benchmarks

Model	MMLU-Pro-Bi			MMLU-Pro-He			GPQA-med			HM-W Exam			HM-T	Average		
	En	Hi	Δ	En	Hi	Δ	En	Hi	Δ	En	Hi	Δ		En	Hi	Δ
Baseline models and HiMed-8B																
GPT-4o	61.9	**55.6**	6.3	57.9	**47.0**	10.9	**41.8**	30.5	11.3	**78.1**	**78.5**	-0.4	62.0	59.9	**54.7**	5.2
UltraMedical-8B	66.8	41.7	25.1	56.1	26.3	29.8	33.3	20.9	12.4	54.7	40.2	14.5	45.1	52.7	34.8	17.9
MedReason-8B	65.8	43.7	22.1	58.7	27.6	31.1	27.3	29.7	-2.4	53.2	41.3	11.9	47.2	51.3	37.9	13.4
HuatuoGPT-o1-8B	71.0	49.1	21.9	59.3	27.6	31.7	37.8	32.9	4.9	59.1	42.8	16.3	56.1	56.8	41.7	15.1
Qwen2.5-7B-Instruct	73.1	41.8	31.3	54.4	20.3	34.1	28.9	19.7	9.2	62.0	41.5	20.5	53.3	54.6	35.3	19.3
Qwen2.5-14B-Instruct	**77.7**	51.9	25.8	**62.1**	26.0	36.1	33.3	26.5	6.8	71.4	46.0	25.4	53.2	**61.1**	40.7	20.4
CURE-Med-7B	65.8	37.9	27.9	44.0	20.8	23.2	27.7	19.7	8.0	65.8	45.8	20.0	44.7	50.8	33.8	17.0
Aya-23-8B	49.1	31.1	18.0	28.0	14.8	13.2	12.4	10.4	2.0	38.6	29.4	9.2	46.6	32.0	26.5	5.5
AyurParam	8.6	7.5	1.1	11.2	12.1	0.9	38.2	43.8	-5.6	27.1	28.6	-1.5	29.8	21.3	24.4	-3.1
LLaMA-3.1-8B-Instruct	66.7	43.5	23.2	52.6	19.7	32.9	31.3	22.5	8.8	52.6	34.3	18.3	51.0	50.8	34.2	16.6
HiMed-8B (ours)	**69.9**	53.8	16.1	**57.8**	32.5	25.3	38.6	36.9	1.7	**58.7**	46.6	12.1	**76.0**	**56.3**	49.2	7.1

HM-W = HiMed-West; HM-T = HiMed-Trad. Bold marks the best score and underline marks the second best within this comparison. HiMed-8B reaches 76.0 on HM-T and reduces the average En-Hi gap from 16.6 to 7.1 points. Ablation highlights appear below.

03 Why Hindi-native reasoning

English-centric medical models lose accuracy in Hindi across both Western and Indian medicine.

Model	MMLU-Pro Hi	Δ	RET Hi	Δ
GPT-4o	50.7	**10.0**	70.0	**18.0**
HuatuoGPT-o1-8B	41.3	**23.4**	26.0	**12.0**
UltraMedical-8B	34.0	**27.3**	28.0	**24.0**

Translation creates a second failure mode

19 of 50 expert-audited GPT-4o translations contained semantic hallucinations. The error below turns an ethnicity into a disease name.

Source. A 15-year-old African-American male with BMI 22 is brought to his physician because of a change in dietary habits.

GPT-4o → Hindi. एक 15 वर्षीय आफ्रिकान द्वीपानोसोमा संक्रमणज रोग पुरुष जिसका BMI 22 है…

Back-translation. “an African trypanosomal-disease male …”

Failure: ethnicity is silently converted into an infectious disease.

Accuracy alone does not select Hindi reasoning

Two responses can receive the same correctness reward even when one gives only a short English rationale. Without a language-form signal, the policy has no incentive to reason in Hindi.

Hindi medical reasoning requires native representations and an explicit training signal for reasoning form.

Cultural grounding cannot be translated away

Indian healthcare combines Western medicine with seven officially recognized Indian systems. Core concepts in Ayurveda, Yoga, Naturopathy, Unani, Siddha, Sowa-Rigpa, and Homoeopathy often lack faithful English equivalents.

The remaining bottlenecks

- Data. Most medical corpora are English-only or omit Indian systems, leaving little native Hindi reasoning supervision.
- Objective. Accuracy rewards the final answer but cannot distinguish a short English rationale from detailed Hindi symptom-to-mechanism reasoning.

HiMed adds native supervision and an objective that rewards Hindi reasoning.

04 Data and training

Two data pipelines with strict decontamination

HiMed-Trad · Indian systems of medicine

- 286,657 corpus instances from classics and examinations
- 6,010 benchmark items across all seven AYUSH systems
- 98.0–99.5% expert-audit acceptance across OCR and generation

HiMed-West · terminology-faithful translation

- 116,859 corpus instances plus 2,254 benchmark items
- 45,902 En–Hi medical terms guide NLLB-200
- 0 hallucinations in audit, versus 19 for GPT-4o

DSR-RL: language scaffolding that fades

I. Language adaptation **II. Reasoning cold start** **III. DSR-RL**

Accuracy reward. Scores correct, reasoned answers with a fine-tuned verifier.

$$R^{acc}_i \in \{0, 0.1, 1\}$$

Language-form reward. Measures the Hindi-token proportion while masking English medical terms for principled code mixing.

$$R^{lan}_i = 1/|y_i| \ \Sigma_t \ \mathbb{I}[y_{i,t} \text{ is a Hindi token}]$$

Combined objective

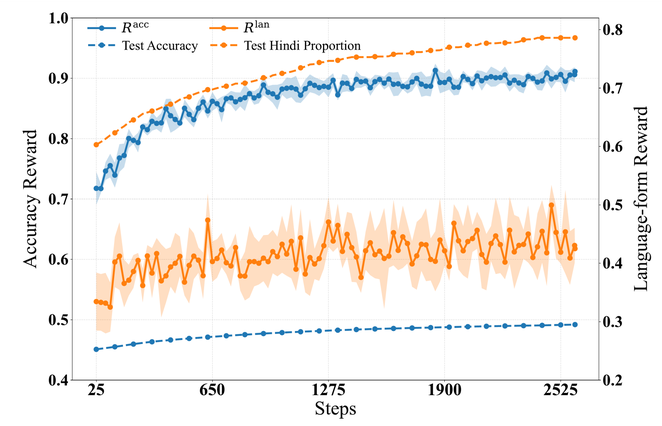
$$\mathcal{J}(\tau) = (1-\lambda(\tau))\ \mathcal{J}^{acc} + \lambda(\tau)\ \mathcal{J}^{lan} - \beta\ \mathrm{KL}$$

Cosine decay emphasizes Hindi form early, then lets accuracy govern the final policy.

Accuracy verifier. A fine-tuned LLaMA-3.2-3B reaches 96.7% agreement with medical experts and 82.3% on adversarial GPT-5 probes.

Scaffolding guides Hindi form early, then decays so task accuracy governs the final policy.

05 Training evidence and evaluation



Reward trends track held-out accuracy and Hindi-token proportion, arguing against reward-only inflation.

67.4% expert preference — non-tie pairs

0.304 open-ended accuracy — base: 0.189

0 language-mixing failures — 88.6% residual errors are knowledge/reasoning

Ablation highlights

Variant	Hi avg	Δ	HM-T
HiMed-8B	**49.2**	**7.1**	**76.0**
full-data SFT	45.4	9.0	74.9
w/o Rlan	46.7	10.0	75.1
binary Rlan	42.7	9.7	72.5
constant λ	43.7	9.2	72.0
linear decay	48.3	7.5	75.5

Open-ended evaluation

Across 500 questions, HiMed-8B improves medical faithfulness from 3.02 to 3.97, reasoning coherence from 3.07 to 3.92, and Hindi nativeness from 2.96 to 3.84. Inter-annotator correlations exceed 0.91.

- The objective matters. Full-data SFT remains below HiMed-8B despite using the same reinforcement-learning data.
- The schedule matters. Dense scaffolding with cosine decay outperforms sparse, fixed, and linear alternatives.

Research and evaluation only. HiMed is not intended for clinical decision-making.

Three takeaways

1 · Data
Hindi-native data covers Western medicine and all seven Indian systems.

2 · Method
Decaying scaffolding guides early Hindi reasoning, then yields to accuracy.

3 · Result
HiMed-8B cuts the En-Hi gap from 16.6 to 7.1 and reaches 76.0% on HiMed-Trad.

Open resources

Code, data, and weights released for research use.

Code
github.com/FreedomIntelligence/HiMed

Data
github.com/FreedomIntelligence/HiMed/tree/main/Data

Weights
huggingface.co/collections/FreedomIntelligence/himed