Big Data and AI: Enhancing Prognostic and Diagnostic Capabilities

Dr. M Sanni Ali, London School of Hygiene and Tropical Medicine,UK
Dr. Seid Muhie Yimam, University of Hamburg, HCDS, DE

01.01.2025

Disclaimer: This talk reflects our computer science and epidemiology perspectives. We are not medical practitioners.

Role of Data in Healthcare Diagnostics and Prognostics

The Role of Data in Modern Healthcare Systems

- **Enhances Care Quality and Efficiency**: Data enables the development of clinical guidelines and predictive models for more personalized, safer, and effective patient care.
- **Informs Strategic Planning**: Data provides insights for healthcare organizations to improve strategic planning and operational quality.
- **Supports Research and Innovation**: Facilitates clinical research and the development of new treatments and medical devices.
- **Influences Policy and Public Health**: Strengthens public health strategies and informs policy-making, contributing to health equity and improved population health.

Exploring Health Data Sources: Access and Utilization

- **Diverse Data Sources**: Includes electronic and non-electronic medical records, disease registries, immunization records, laboratory data, and national surveys.
- **Comprehensive Data Utilization**: Essential for healthcare reform, clinical evaluation, and public health planning.
- **Improved Health Outcomes**: Access to detailed health data allows for better analysis of healthcare delivery and patient outcomes.
- **Need for Infrastructure and Skills**: Effective utilization requires investment in data governance, technical infrastructure, and workforce training.

Data Sources

- Electronic and Non-Electronic Medical Records.
- Disease registries such as Cancer.
- Disease notification records - EPHI.
- Birth and death registries: National Statistics Agency.
- Immunisation records.
- Laboratory records.
- National health surveys and research data.

https://www.linkedin.com/pulse/transforming-healthcare-power-data-driven-insights-kevin-sila/

Benefits of Big Data Analytics for Healthcare

- Individual patient care
- Predicting the cost and risks of treatment
- Prevention of mass diseases and preventive care
- New therapy and drug discovery
- Early detection of diseases
- Real-time alert
- Optimized hospital operation
- Preventing unnecessary emergency room visits
- Fewer medical errors and more accurate treatment
- Modeling the spread of disease
- Improved HR management
- Identifying and managing high-risk patients
- Prevention of suicide and self-harm
- Better customer service
- Reducing the cost of medical care

Here are the trends emerging from efforts to bring big data to healthcare:

01

Driving from models of emergency and ad hoc care to value-based care with hospital analytics.

02

Empower you to capitalize on extensive collections of medical data with clinical analytics.

03

Reliable patient analytics, using data from a combination of sources to find it and help solve complex health problems.

Analytics Implementation Stages

Stage 1	Stage 2	Stage 3	Stage 4	Stage 5
Developing a Big Data strategy	Identify Big Data Sources	Developing methods for accessing, managing, and storing	Analyzing Big Data	Make well-informed decisions

Enhancing Diagnostics and Prognostics: Case Studies

Diagnostic epidemiology

- **Example:** Ruling out deep venous thrombosis in primary care: a simple diagnostic algorithm using D-dimer testing.
  (R Oudega, KGM Moons, AW Hoes, Thromb Haemost 2005).
- In primary care, physicians have to decide which patients have to be referred for further diagnostic work-up.
- In 2005, only in 20% - 30% of the referred patients, diagnosis of DVT is confirmed.
- This puts a burden on both patients and health care budget.

- Whether the diagnostic work-up and referral of patients suspected of DVT in primary care could be more efficient.

$$\text{DVT}+ = f(D_1, D_2, D_3, \ldots)$$

- Cross-sectional study, 1295 adult patients visiting the primary care physicians in three hospitals in the Netherlands in whom DVT is suspected by the physician on clinical grounds.

- Suspicion of DVT was based on the presence of at least one of the following symptoms or signs of lower extremities: swelling, redness, and/or pain in the legs.

- Following **history**, findings recorded as potential diagnostic information included: DVT(+), family history of DVT, history of any malignancy(active cancer in the previous 6 months, immobilization($>$ 3 days), recent surgery (past 4 weeks), leg trauma (past 4 weeks), pain when walking, presence of the duration of the three main symptoms(redness, swelling, pain in the leg).

- **Physical examination** items included tenderness along deep vein system in calf or thigh, distention of collateral veins in the symptomatic leg, pitting edema, 3cm or more difference in circumference of the calves.

Diagnostic variables	Total n=1295 %	DVT present n=289 %	DVT absent n=1006 %	OR (95% CI)
Patient history:				
age (years)	60.0 (17.6)[1]	62.0 (16.8)[1]	59.4 (17.8)[1]	1.01 (1.00 – 2.02)[2]
gender + OC use				
males	36	47	33	1.95 (1.47 – 2.57)
females using OC	10	10	10	1.37 (0.87 – 2.17)
females not using OC	54	43	57	-
gender + HRT use				
males	36	47	33	1.86 (1.42 – 2.43)
females using HRT	2	2	2	1.32 (0.48 – 3.63)
females not using HRT	62	51	66	-
previous DVT	24	21	25	0.82 (0.60 – 1.12)
family history of DVT	23	20	24	0.79 (0.57 – 1.09)
presence of malignancy	6	12	5	2.72 (1.71 – 4.32)
immobilization	14	13	14	0.90 (0.61 – 1.33)
recent surgery	14	19	13	1.59 (1.12 – 2.26)
absence of leg trauma	85	89	84	1.58 (1.05 – 2.36)
pain when walking	81	84	80	1.30 (0.92 – 1.84)
days of symptoms	7.9 (7.6)[1]	6.9 (6.7)[1]	8.2 (7.8)[1]	0.98 (0.96 – 0.99)[3]
Physical examination:				
vein distension	20	28	17	1.88 (1.39 – 2.55)
deep vein system tenderness	71	72	71	1.04 (0.78 – 1.39)
swelling whole leg	45	57	42	1.84 (1.41 – 2.39)
calf difference ≥ 3cm	43	67	36	3.63 (2.75 – 4.79)
D-dimer abnormal				
VIDAS n= 918	78	99	72	38.2 (9.40 – 155.3)
Tinaquant n= 377	65	98	54	37.3 (9.00 – 154.8)
Combined assays	74	99	66	35.7 (13.3 - 100.0)

Table 2: Independent diagnostic indicators of DVT. The final multivariate model, the figures are estimated after model validation and adjustment for over-fitting.

Diagnostic variables	Odds ratio	Regression coefficient*	p-value	Points for the rule
Male gender	1.80 (1.36 – 2.16)	0.59	<0.001	1
Oral contraceptive use	2.12 (1.32 – 3.35)	0.75	0.002	1
Presence of malignancy	1.52 (1.05 – 2.44)	0.42	0.082	1
Recent surgery	1.46 (1.02 – 2.09)	0.38	0.044	1
Absence of leg trauma	1.82 (1.25 – 2.66)	0.60	0.002	1
Vein distension	1.62 (1.19 – 2.20)	0.48	0.002	1
Calf difference ≥ 3 cm	3.10 (2.36 – 4.06)	1.13	<0.001	2
D-dimer abnormal	20.3 (8.25 – 49.9)	3.01	<0.001	6
Constant		–5.47		

DVT= deep vein thrombosis; *=natural logarithm of the odds ratio; D-dimer abnormal for VIDAS ≥ 500 ng/ml and Tinaquant ≥ 400 ng/ml. Probability of DVT as estimated by the final model =1/(1+exp-(-5·47 + 0·59*male gender + 0·75*OC use + 0·42*presence of malignancy + 0·38*recent surgery + 0·60*absence of leg trauma + 0·48*vein distension + 1·13*calf difference ≥ 3cm + 3·01*abnormal D-dimer)).

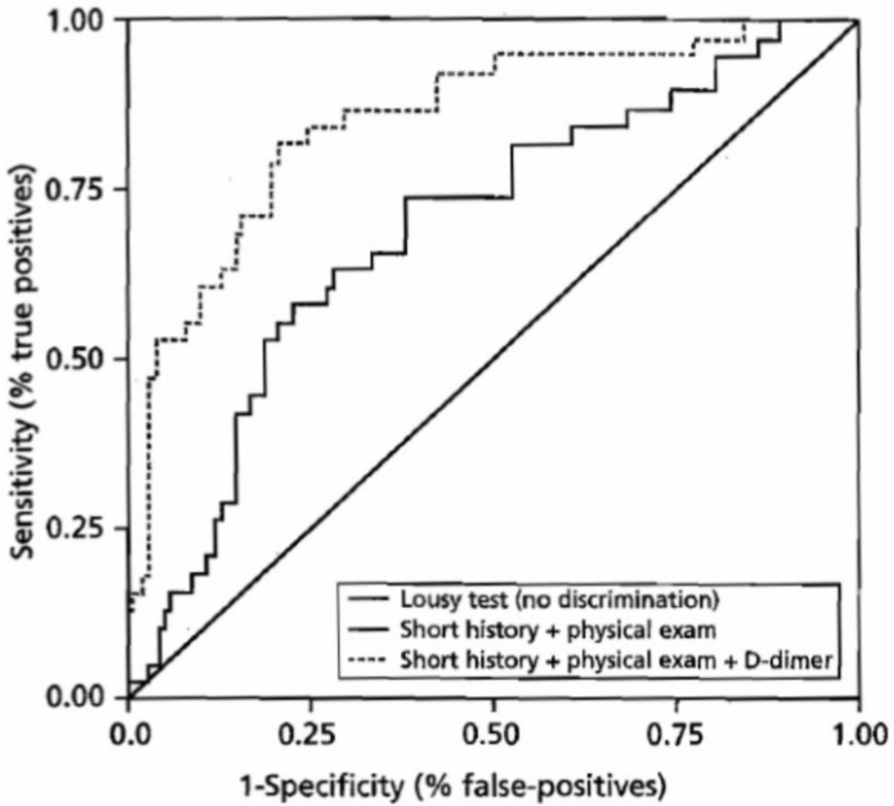


FIGURE 3.3 Example of an ROC curve of the reduced multivariable logistic regression model, including the same six determinants as in Figure 3.2. The ROC area of the "reduced history + physical model" (red) was 0.70 (95% confidence interval [CI], 0.66–0.74) and of the same model added with the D-dimer assay (green) 0.84 (95% CI, 0.80–0.88).

$1*\text{male gender} + 1*\text{OC use} + 1*\text{presence of malignancy} + 1*\text{recent surgery} + 1*\text{absence of trauma} + 1*\text{vein distension} + 2*\text{calf difference} \geq 3\text{cm} + 6*\text{abnormal D-dimer test}.$

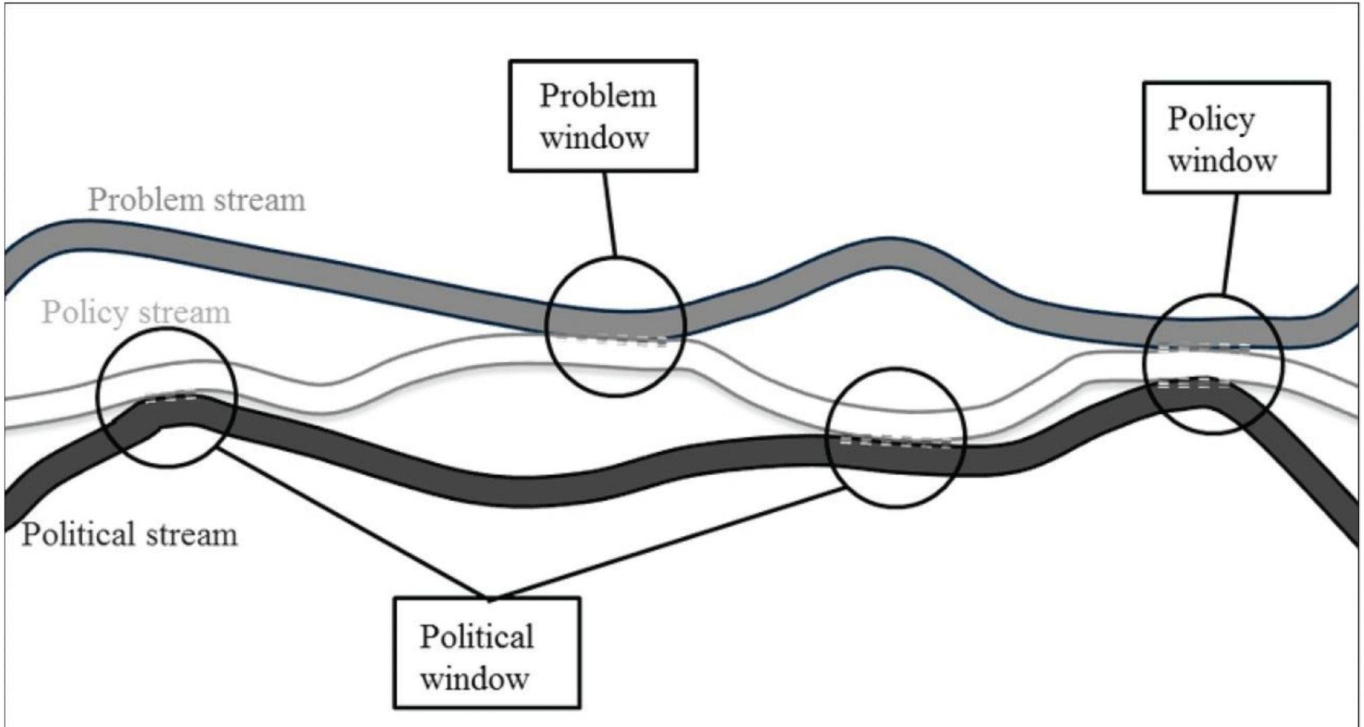
Table 4: Prevalence of DVT across four score (risk) categories.

Probability or risk Category	number of patients n (%)[1]	DVT present n (%)[2]	DVT absent n (%)[3]
Very low (0–3)	293 (23)	2 (0.7)	291 (99.3)
Low (4–5)	66 (5)	3 (4.5)	63 (95.5)
Moderate (7–9)	663 (51)	144 (21.7)	519 (78.3)
High (10–13)	273 (21)	140 (51.3)	133 (48.7)

1=proportion of all (1295) patients; 2=proportion of presence of DVT within risk category; 3=proportion of absence of DVT within risk category.

Integrating Health in All Policies: A Comprehensive Approach

Figure: Schematic representation of Kingdon's non-linear framework for policy-making



Enhancing Medication Policies Through Data Insights

Protelos and Osseor

Share

CURRENT STATU
European Commis

Procedure started — Under evaluation — PRAC recommendation — CHMP opinion — European Commission final decision

Table of contents

Overview

Protelos/Osseor to remain available but with further restrictions

The European Medicines Agency has concluded its review of Protelos/Osseor and has recommended further restricting the use of the medicine to patients who cannot be treated with other medicines approved for osteoporosis. In addition these patients should continue to be evaluated regularly by their doctor and treatment should be stopped if patients develop heart or circulatory problems, such as uncontrolled high bloo pressure or angina. As recommended in a previous review, patients who have a history of certain heart or circulatory problems, such as stroke and heart attack, must not use the medicine.

These final recommendations from the Agency's Committee for Medicinal Products for Human Use (CHMP) come after initial advice from the Pharmacovigilance Risk Assessment Committee (PRAC) to suspend the medicine due to its cardiovascular risk.

Springer Link

Original Article | Published: 06 November 2019

Impact of risk minimisation measures on the use of strontium ranelate in Europe: a multi-national cohort study in 5 EU countries by the EU-ADR Alliance

K. Berencsi, A. Sami, M.S. Ali, K. Marinier, N. Deltour, S. Perez-Gutthann, L. Pedersen, P. Rijnbeek, J. Van der Lei, F. Lapi, M. Simonetti, C. Reyes, M.C.J.M. Sturkenboom & D. Prieto-Alhambra

Osteoporosis International **31**, 721–755 (2020) | Cite this article

Original Article | Published: 05 August 2020

Comparative cardiovascular safety of strontium ranelate and bisphosphonates: a multi-database study in 5 EU countries by the EU-ADR Alliance

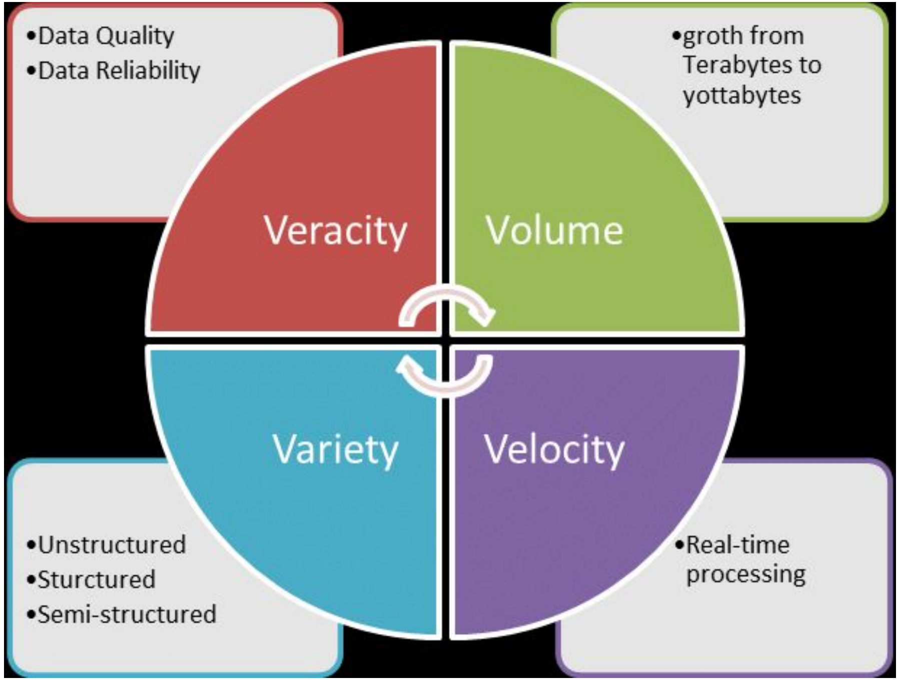
M.S. Ali, K. Berencsi, K. Marinier, N. Deltour, S. Perez-Guthann, L. Pedersen, P. Rijnbeek, F. Lapi, M. Simonetti, C. Reyes, J. Van der Lei, M. Sturkenboom & D. Prieto-Alhambra

Osteoporosis International **31**, 2425–2438 (2020) | Cite this article

Harnessing Big Data for Diagnostic and Prognostic Innovation

- **AI and Analytics**: Utilizes advanced algorithms to interpret vast data for predictive analytics and automated decision-making.
- **Personalized Medicine**: Enables tailored healthcare strategies by analyzing individual genetic and health data.
- **Operational Improvements**: Streamlines processes and optimizes resource use through data-driven insights.
- **Research and Development**: Accelerates the discovery of new treatments and medical advancements.
- **Data Security and Governance**: Necessitates robust measures to protect patient privacy and ensure data integrity.

Dissecting Big Data: Veracity, Volume, Velocity, and Variety



https://www.researchgate.net/profile/Mohammad-Ashraf-Ottom/publication/313783767/figure/fig3/AS:648219604287493@1531559005823/Big-data-in-healthcare-4Vs.png

- **Integration Challenges**: Managing and integrating data from disparate sources to form a cohesive dataset for analysis.
- **Impact on Decision-Making**: The 3Vs enable more informed and agile clinical and operational decisions in healthcare.

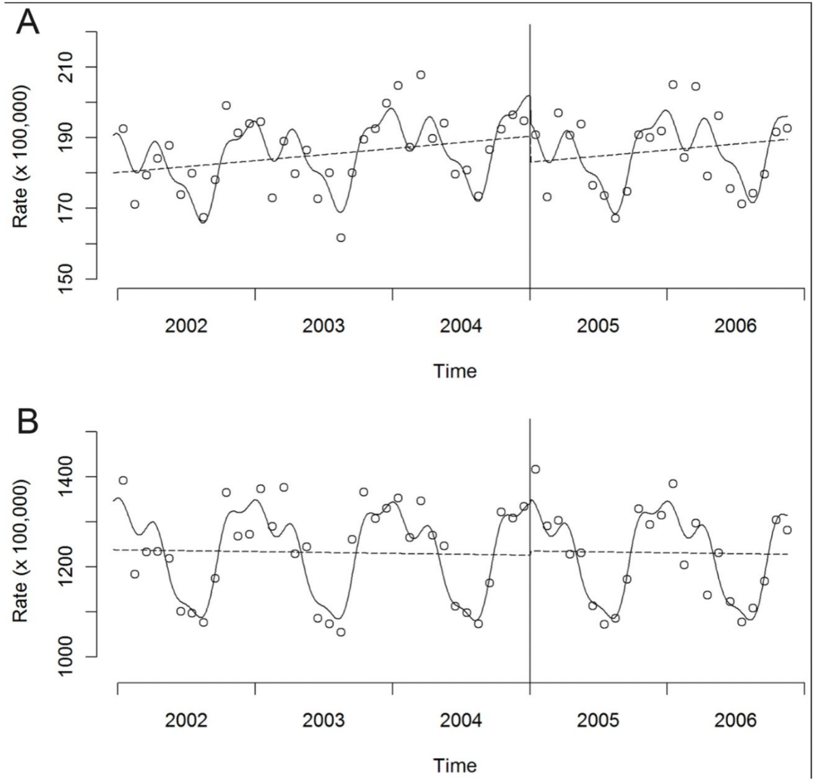
Statistical Approaches to Enhance Diagnostic and Prognostic Models

- Machine learning methods
- Deep learning
- Image data processing
- Data visualization (Dashboard)

Utilizing Evidence for Effective Policy Decision-Making

Case Study: Using Data for Successful Smoking Ban Policies

Hospital admissions for ACEs in Italy during 2002–2006.



Italy smoking ban: Sicily

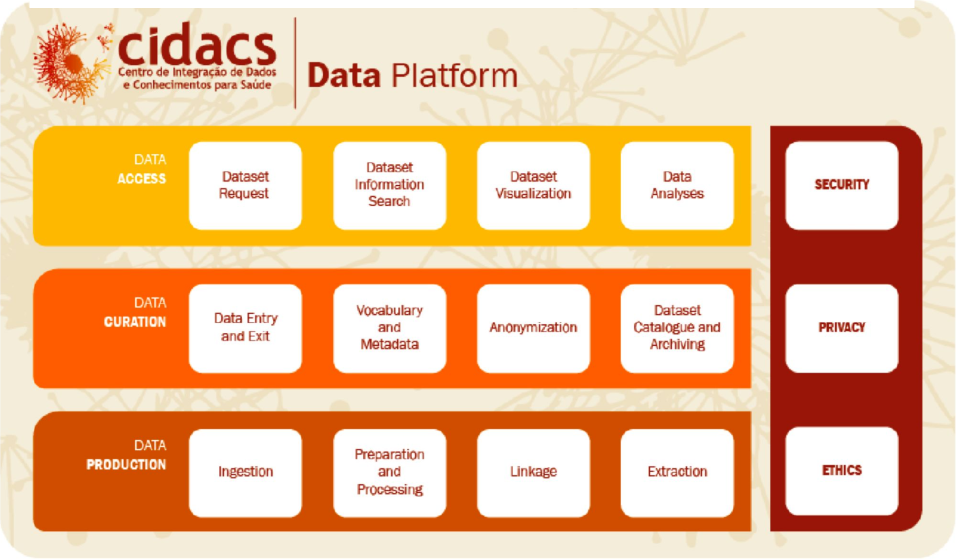
Evaluating Socio-Economic Health Programs Using Data

Cohort Profile: The 100 Million Brazilian Cohort

Mauricio L Barreto, Maria Yury Ichihara, Julia M Pescarini, M Sanni Ali, Gabriela L Borges, Rosemeire L Fiaccone, Rita de Cássia Ribeiro-Silva, Carlos A Teles, Daniela Almeida, Samila Sena ... Show more

International Journal of Epidemiology, Volume 51, Issue 2, April 2022, Pages e27–e38,
https://doi.org/10.1093/ije/dyab213

Published: 18 December 2021 **Article history ▾**



Barbosa *et al. BMC Med Inform Decis Mak* (2020) 20:289
https://doi.org/10.1186/s12911-020-01285-w

BMC Medical Informatics and Decision Making

RESEARCH ARTICLE Open Access

CIDACS-RL: a novel indexing search and scoring-based record linkage system for huge datasets with high accuracy and scalability

George C. G. Barbosa[1*], M. Sanni Ali[1,2,3], Bruno Araujo[1], Sandra Reis[1], Samila Sena[1], Maria Y. T. Ichihara[1], Julia Pescarini[1], Rosemeire L. Fiaccone[1,4], Leila D. Amorim[1,4], Robespierre Pita[1], Marcos E. Barreto[1,6,7], Liam Smeeth[2] and Mauricio L. Barreto[1,5]

Open access Protocol

BMJ Open

Evaluating the impact of the Bolsa Familia conditional cash transfer program on premature cardiovascular and all-cause mortality using the 100 million Brazilian cohort: a natural experiment study protocol

Julia M Pescarini ,[1,2] Peter Craig,[3] Mirjam Allik,[3] Leila Amorim,[4] Sanni Ali,[1,5] Liam Smeeth,[5,6] Mauricio L Barreto,[1,7] Alastair H Leyland,[3] Estela M L Aquino,[1,7] Srinivasa Vittal Katikireddi [3]

Abbreviation	Year	Registers
CadUnico	2003	Individuals and their socio-economic characteristic applying for social benefits.
BFP	2003	Individuals receiving BF payments.
SINASC	1990	All births in Brazil including the type of pregnancy and delivery.
SIM	1975	All deaths in Brazil including ICD-10 cause of death.
SINAN	1993	Diseases of compulsory notification using ICD-10 codes.
SIH-SUS	1993	Patient admissions in the network of public hospitals under SUS.
SIA-SUS	1995	Outpatient visits by SUS.
APAC-SIA	1996	High-cost ambulatory procedures and high-cost medicines.
RHC	1967	Cancer patients in (public or private) hospitals responsible for oncology care.
RCBP	1967	Cancer patients in centers located mostly in major cities.
SISMAMA	2004	Information about breast and gynaecological cancer screening.
SI-PNI	1973	Dispensed immunobiologicals.
SIAB-SUS	1998	Home visits, and medical and nursing care performed in households and health unit
SISLAB-GAL	2008	Laboratory test including cases of Compulsory Notification.
NOTIVISA	2008	Spontaneous reports of suspected cases of Adverse Drug Events.

Open access Protocol

BMJ Open

Evaluating the health effect of a Social Housing programme, Minha Casa Minha Vida, using the 100 million Brazilian Cohort: a natural experiment study protocol

Andrêa J F Ferreira ,[1,2] Julia Pescarini ,[3] Mauro Sanchez,[4] Renzo Joel Flores-Ortiz,[5] Camila Silveira Teixeira,[1] Rosemeire Fiaccone,[6] Maria Yury Ichihara,[1] Rodrigo Oliveira,[7] Estela M L Aquino,[1] Liam Smeeth,[8] Peter Craig,[9] Sanni Ali,[10] Alastair H Leyland,[9] Mauricio L Barreto,[1] Rita de Cássia Ribeiro,[1] Srinivasa Vittal Katikireddi [11]

RESEARCH ARTICLE

Conditional cash transfer program and child mortality: A cross-sectional analysis nested within the 100 Million Brazilian Cohort

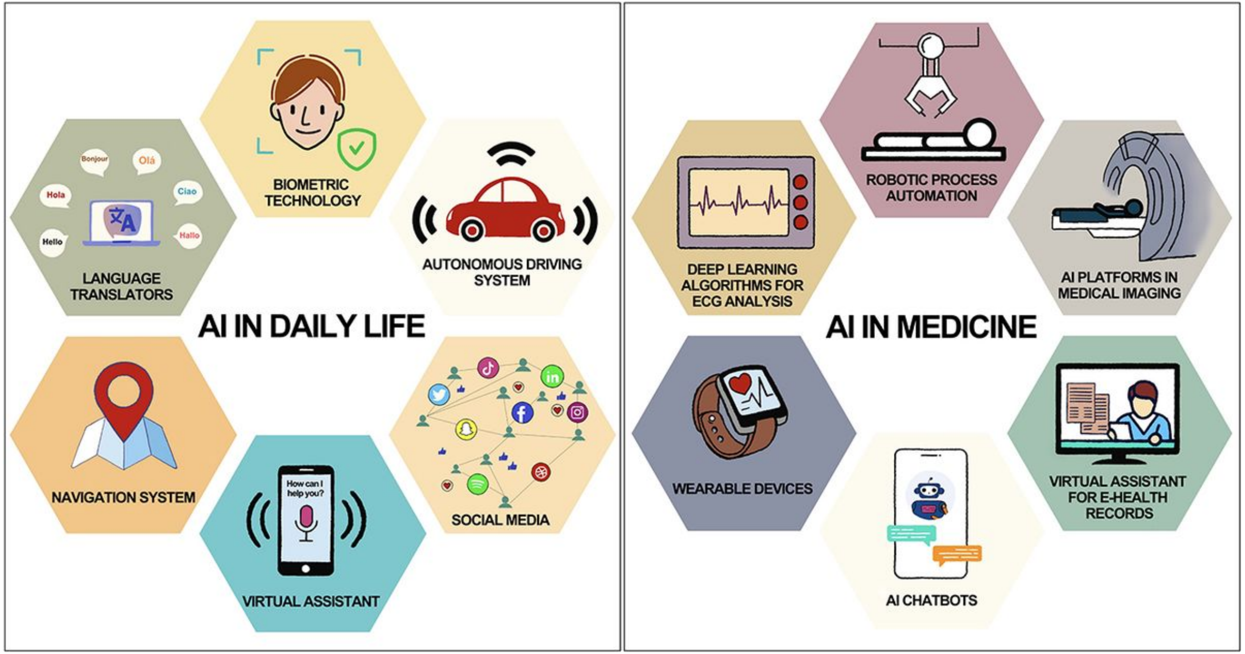
Dandara Ramos[1,2]*, Nívea B. da Silva[1,3], Maria Yury Ichihara[1,2], Rosemeire L. Fiaccone[1,3], Daniela Almeida[1,4], Samila Sena[1], Poliana Rebouças[1,2], Elzo Pereira Pinto Júnior[1], Enny S. Paixão[1,5], Sanni Ali[1,5], Laura C. Rodrigues[1,5], Maurício L. Barreto[1,2]

1 Center for Data and Knowledge Integration for Health (CIDACS), Fundação Oswaldo Cruz, Salvador, Bahia, Brazil, **2** Institute of Collective Health, Federal University of Bahia, Salvador, Bahia, Brazil, **3** Statistics Department, Institute of Mathematics and Statistics, Federal University of Bahia, Salvador, Bahia, Brazil, **4** Computer Science Department, Institute of Mathematics and Statistics, Federal University of Bahia, Salvador, Bahia, Brazil, **5** Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, United Kingdom

AI and LLM Applications in Prognostics and Diagnostics

Artificial Intelligence in Medical Diagnosis

- AI is used to analyze vast amount of data quickly and accurately.
  - assisting healthcare providers in making more informed decisions.



https://images-provider.frontiersin.org/api/ipx/w=1200&f=png/https://www.frontiersin.org/files/Articles/1227091/frai-06-1227091-HTML/image_m/frai-06-1227091-g001.jpg

https://www.spectral-ai.com/blog/artificial-intelligence-in-medical-diagnosis-how-medical-diagnostics-are-improving-through-ai/

Role of AI in Diagnosis

- Enhanced **Accuracy**: AI algorithms improve diagnostic accuracy by analyzing complex medical data, reducing **human error**.
- **Early Detection**: Machine learning models can identify early **signs of diseases** such as cancer or heart disease, allowing for **timely intervention**.
- **Personalized** Medicine: AI tailors treatments based on **individual patient** data, leading to more effective and **personalized care plans**.
- **Efficiency**: Automated systems **speed up** the **diagnostic process**, freeing up healthcare providers to focus on patient care.

Benefits of AI in Healthcare

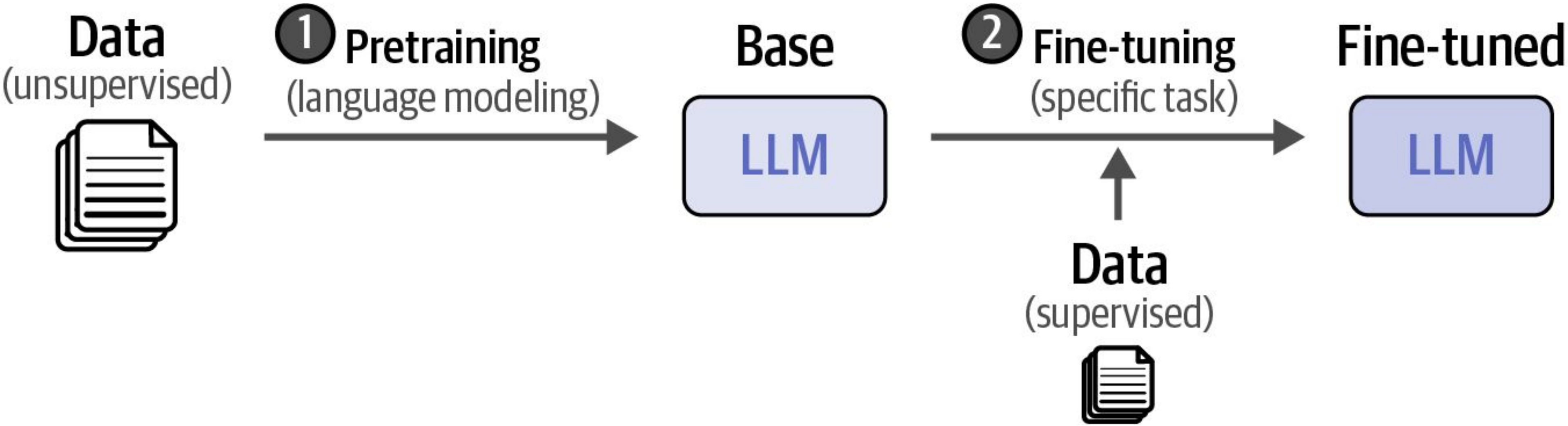
- **Data Analysis**: AI processes large datasets from electronic health records (EHRs), providing insights that are difficult to achieve manually.
- **Imaging**: Advanced AI tools enhance the interpretation of medical images, aiding radiologists in identifying abnormalities.
- **Predictive Analytics**: Predictive models forecast disease progression, helping in preventive care and better resource allocation.
- **Clinical Decision Support**: AI systems provide evidence-based recommendations, supporting clinicians in making more informed decisions.

DeepView® technology

Introduction to Large Language Models (LLMs)

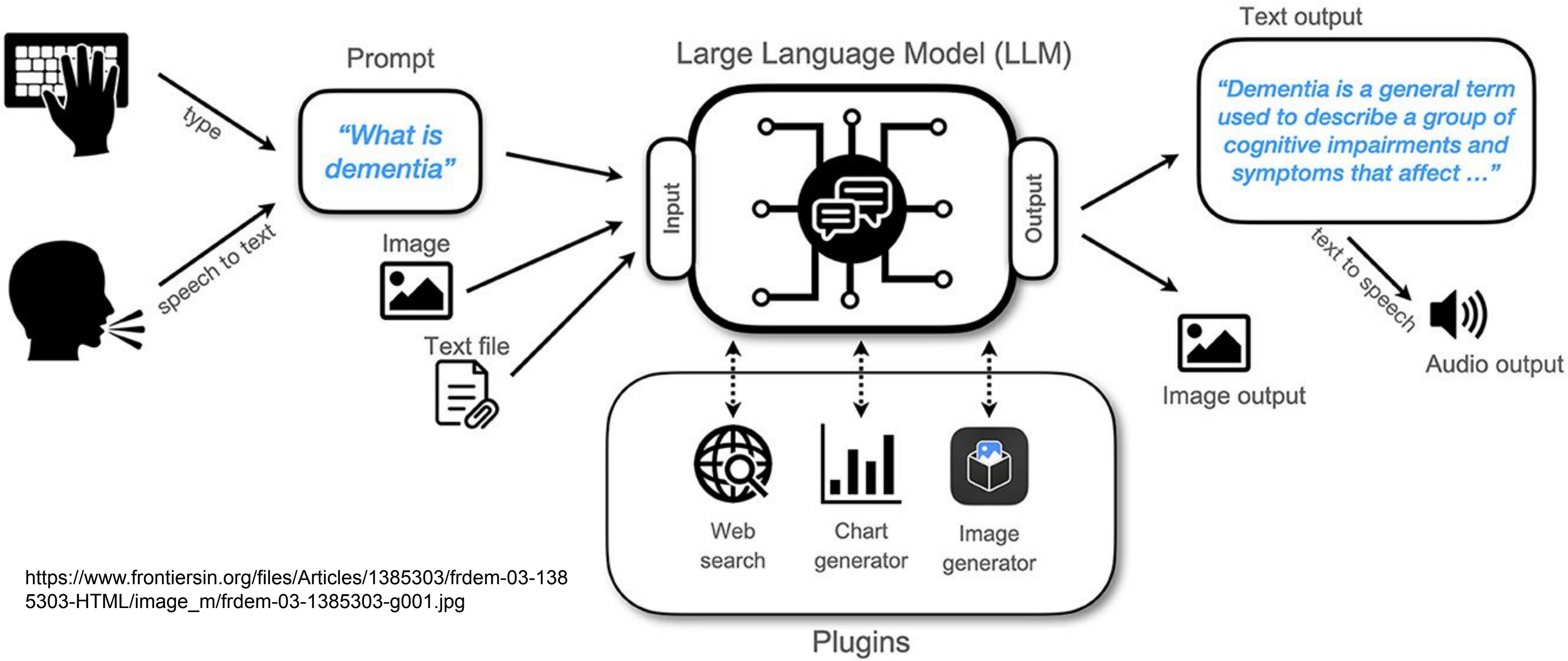
- **Language Models**: Predict and generate text based on the probability of sequences within a text. Useful for tasks like autocomplete, text generation, and translation.
- **Large Language Models** (LLMs): Advanced models with vast parameters and datasets. Capable of handling complex tasks, such as summarization and question answering.
- **Transformers**: Key architecture for LLMs, using attention mechanisms to improve processing by focusing on significant input aspects.
- **Considerations**: LLMs come with high cost, potential bias, and ethical concerns. They demand significant resources for training and special infrastructure for deployment.

Large Language Models (LLM) Pipeline



https://mirascope.com/assets/blog/llm-pipeline/stages_of_rag_pipeline.png

LLMs in Healthcare



https://www.frontiersin.org/files/Articles/1385303/frdem-03-1385303-HTML/image_m/frdem-03-1385303-g001.jpg

LLMs in Healthcare

Model	Developer	Year of Release	Parameters	Multimodal	Primary Use Case	Availability
MedLM	Google	2023	340B	✅	Medical question answering	Closed-source
RadOnc GPT	Meta	2023	70B	❌	Radiology image analysis	Open-source
MedAlpaca	Technical University of Munich	2023	13B	❌	Clinical data analysis	Open-source
GatorTron	NVIDIA	2021	3.9B	❌	Medical NLP	Closed-source
BioMedLM	Stanford University	2022	2.7B	❌	Biomedical research	Open-Source

Billing & Coding: Automates accurate billing and coding, reducing errors.
Appointments: Chatbots efficiently schedule appointments based on availability.
Report Generation: Drafts health status reports from patient data.

Virtual Assistant: Handles inquiries and scheduling, offering triage support.
Language Interpretation: Bridges language gaps during teleconsultations.
Emotional Dissection: Detects patient emotions for better support.

Diagnostic: AMIE outperforms human accuracy with advanced medical training.
Patient Interaction: Provides empathetic communication and critical insights.
Multi Agent Training: Enhances interaction precision and empathy through simulations.

Biomedical Research: LLMs accelerate discovery and validation of new biological models.
Drug Research: LLMs generate research hypotheses to guide therapeutic development.

LLM and Prompting for Healthcare

- **LLM prompts** guide models to produce specific outcomes by providing structured input with instructions and context.
- **Challenges Mitigation**:
  - Well-crafted prompts minimize hallucinations and biases by focusing responses.
- **Clinical Examples**:
  - Summarization: "***Summarize patient's diagnosis and treatment plan from the Aug 5th appointment***."
  - Information Extraction: "***List key symptoms and medications from these clinical notes***."
  - Plain Language Translation: "***Translate clinical notes for patient understanding, ensuring accuracy***."

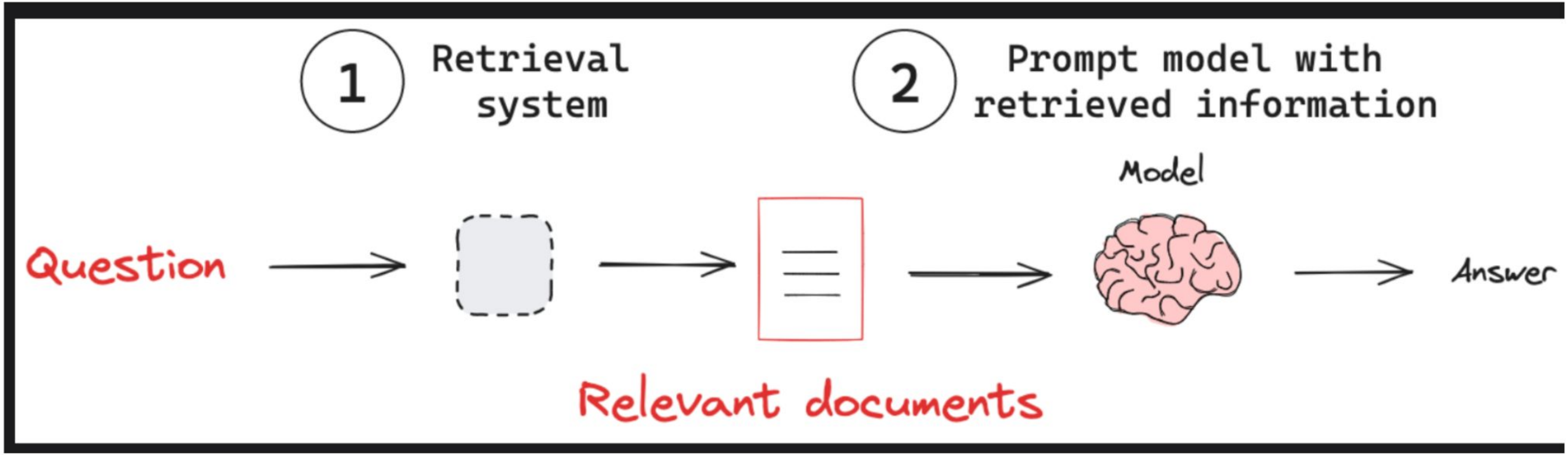
Shah, K., Xu, A. Y., Sharma, Y., Daher, M., McDonald, C., Diebo, B. G., & Daniels, A. H. (2024). Large language model prompting techniques for advancement in clinical medicine. *Journal of Clinical Medicine, 13*(17), 5101. https://doi.org/10.3390/jcm13175101

Practical examples

- Examples,
  - Llama 3 and Open AI GPT4-o-mini: [Simple example]()

```
def get_completion_llama(prompt, model_pipeline=llama3):
    messages = [{"role": "user", "content": prompt}]
    response = model_pipeline(
        messages,
        max_new_tokens=2000
    )
    return response[0]["generated_text"][-1]['content']
```

Retrieval Augmented Generation, LLM and Healthcare



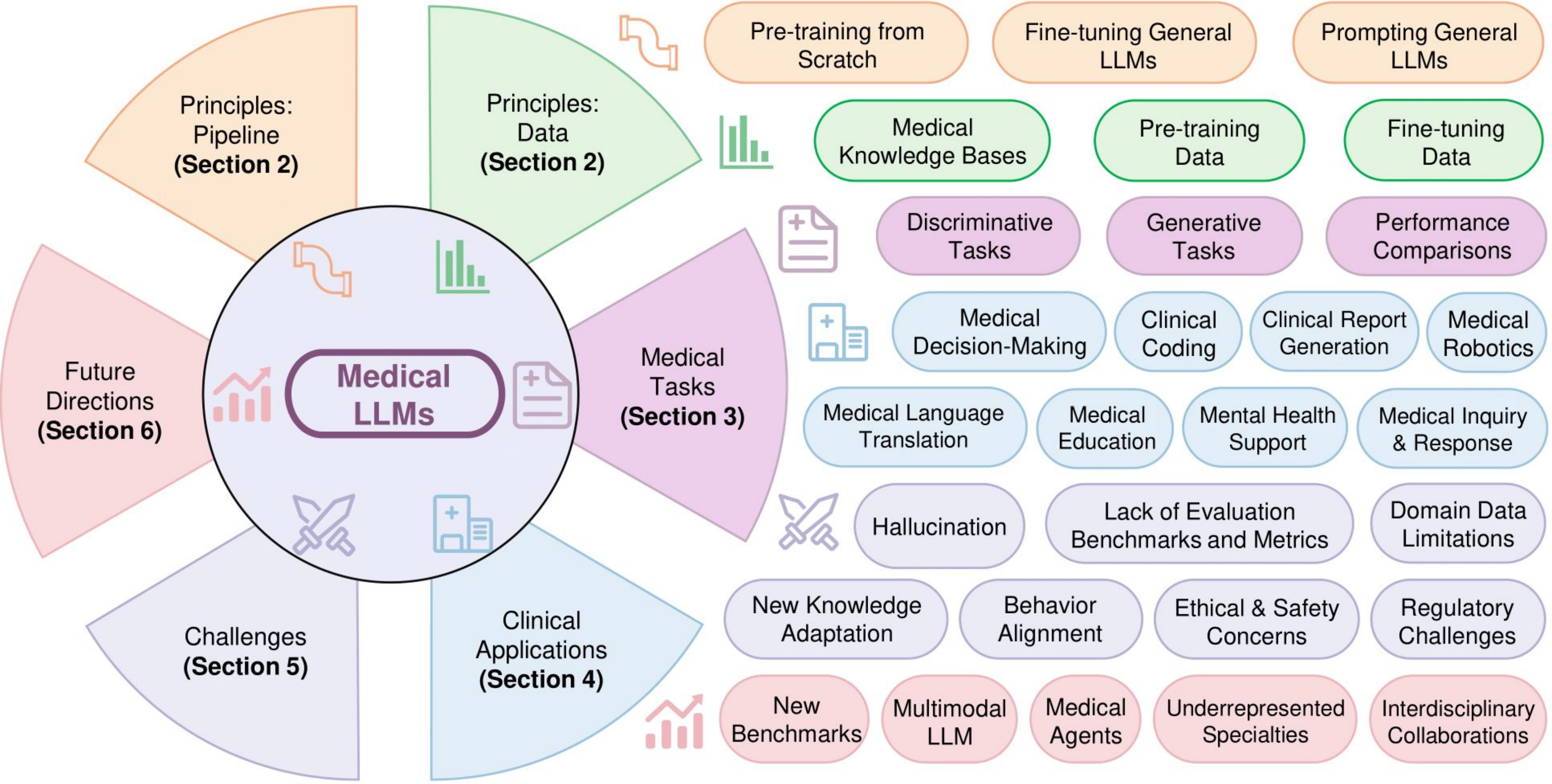
- Llama 3 and Open AI GPT4-o-mini: [Simple example]

A Survey of Large Language Models in Medicine: Progress, Application, and Challenge

- Overview:
    - Comprehensive analysis of large language models (LLMs) in medicine, addressing principles, applications, and challenges.
- Key Questions:
    - How are medical LLMs constructed?
    - What metrics assess their downstream performance?
    - How can they be applied in clinical practice?
    - What challenges do they face, and how can they be optimized?
- Objective:
    - Provide insights into opportunities and challenges of medical LLMs and serve as a practical guide for their effective construction and utilization.

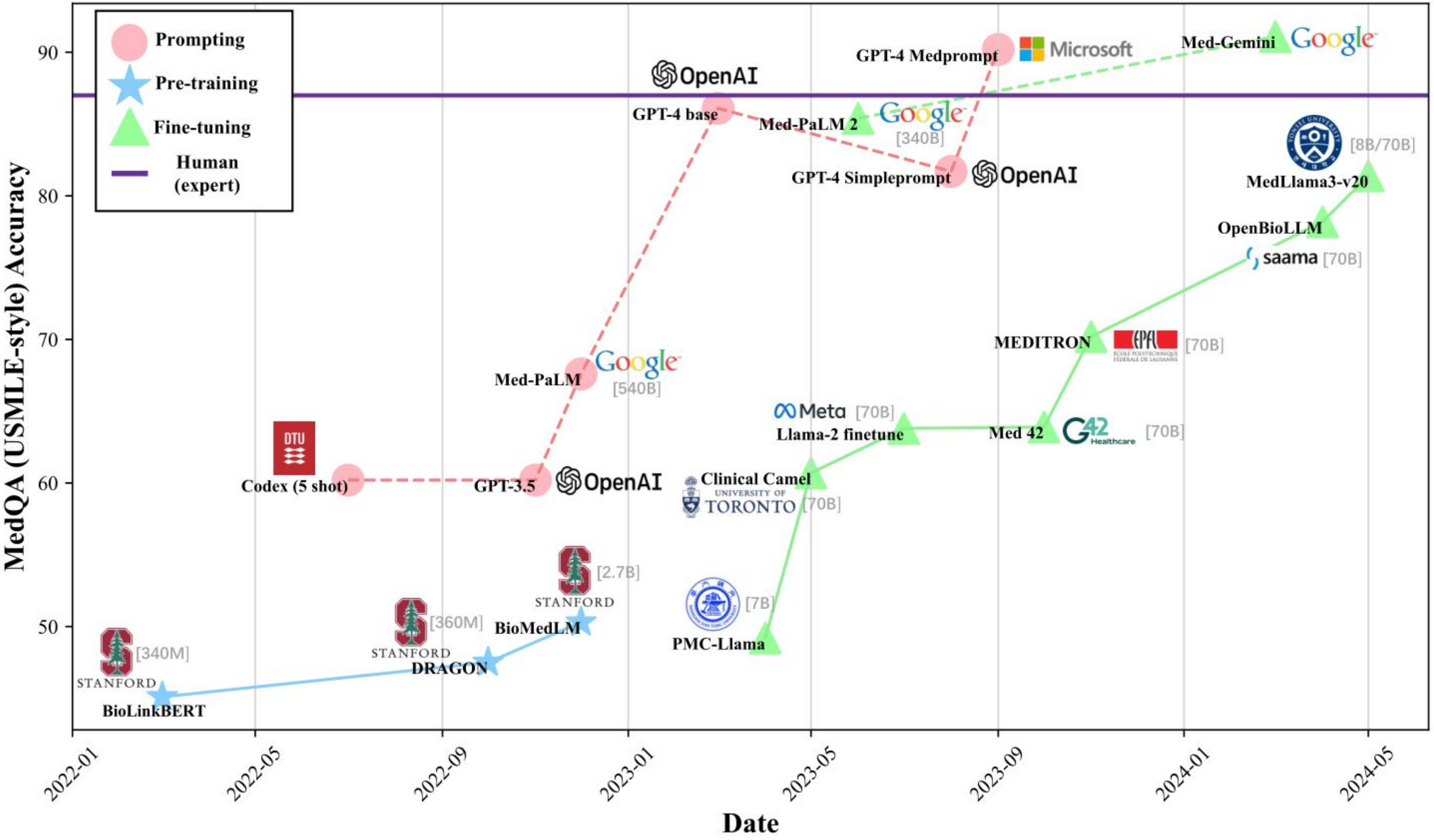
https://github.com/AI-in-Health/MedLLMsPracticalGuide

Zhou, H., Liu, F., Gu, B., Zou, X., Huang, J., Wu, J., Li, Y., Chen, S. S., Zhou, P., Liu, J., Hua, Y., Mao, C., Wu, X., Zheng, Y., Clifton, L., Li, Z., Luo, J., & Clifton, D. A. (2023). *A Survey of Large Language Models in Medicine: Progress, Application, and Challenge*. arXiv preprint arXiv:2311.05112.

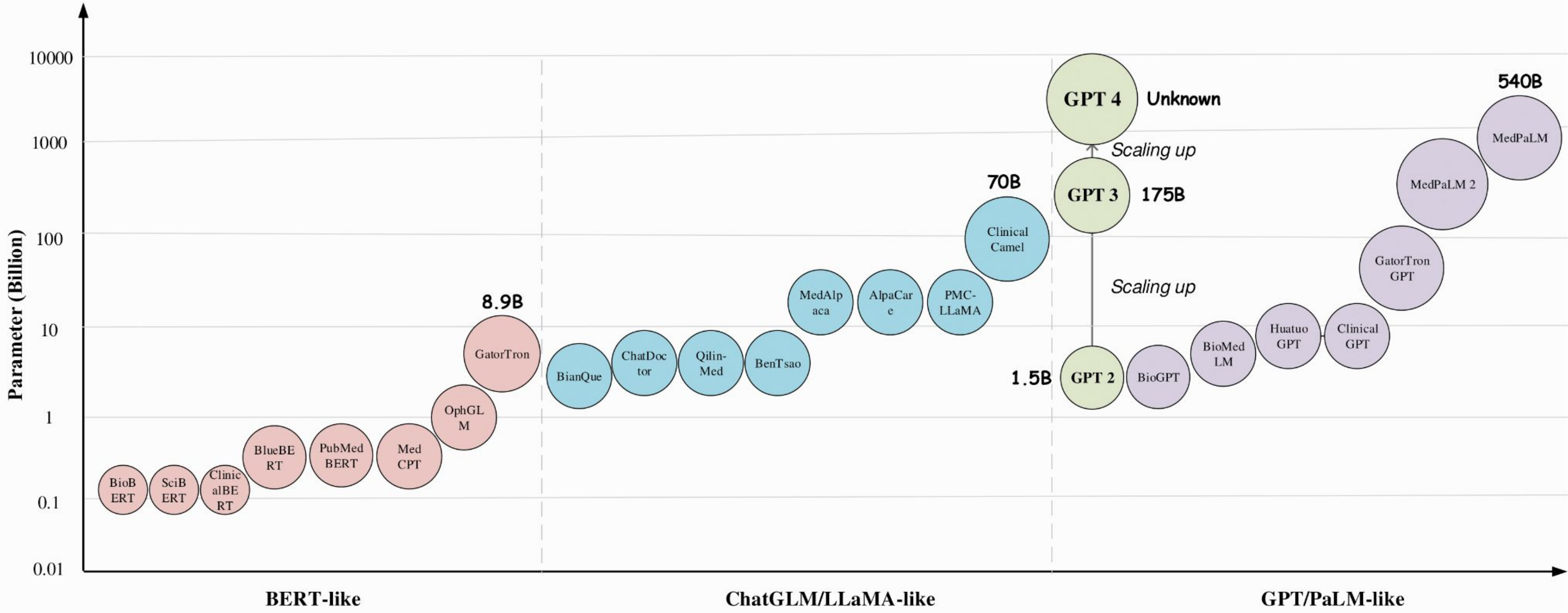

Principles: Pipeline (Section 2)
Principles: Data (Section 2)
Medical Tasks (Section 3)
Clinical Applications (Section 4)
Challenges (Section 5)
Future Directions (Section 6)
Medical LLMs
Pre-training from Scratch
Fine-tuning General LLMs
Prompting General LLMs
Medical Knowledge Bases
Pre-training Data
Fine-tuning Data
Discriminative Tasks
Generative Tasks
Performance Comparisons
Medical Decision-Making
Clinical Coding
Clinical Report Generation
Medical Robotics
Medical Language Translation
Medical Education
Mental Health Support
Medical Inquiry & Response
Hallucination
Lack of Evaluation Benchmarks and Metrics
Domain Data Limitations
New Knowledge Adaptation
Behavior Alignment
Ethical & Safety Concerns
Regulatory Challenges
New Benchmarks
Multimodal LLM
Medical Agents
Underrepresented Specialties
Interdisciplinary Collaborations

What are the Goals of the Medical LLM?

Goal 1: Surpassing Human-Level Expertise.



Goal 2: Emergent Properties of Medical LLM with the Model Size Scaling Up.



Article | Open access | Published: 05 July 2024

Pre-trained multimodal large language model enhances dermatological diagnosis using SkinGPT-4

https://www.nature.com/articles/s41467-024-50043-3

Juexiao Zhou, Xiaonan He, Liyuan Sun, Jiannan Xu, Xiuying Chen, Yuetan Chu, Longxi Zhou, Xingyu Liao, Bin Zhang, Shawn Afvari & Xin Gao

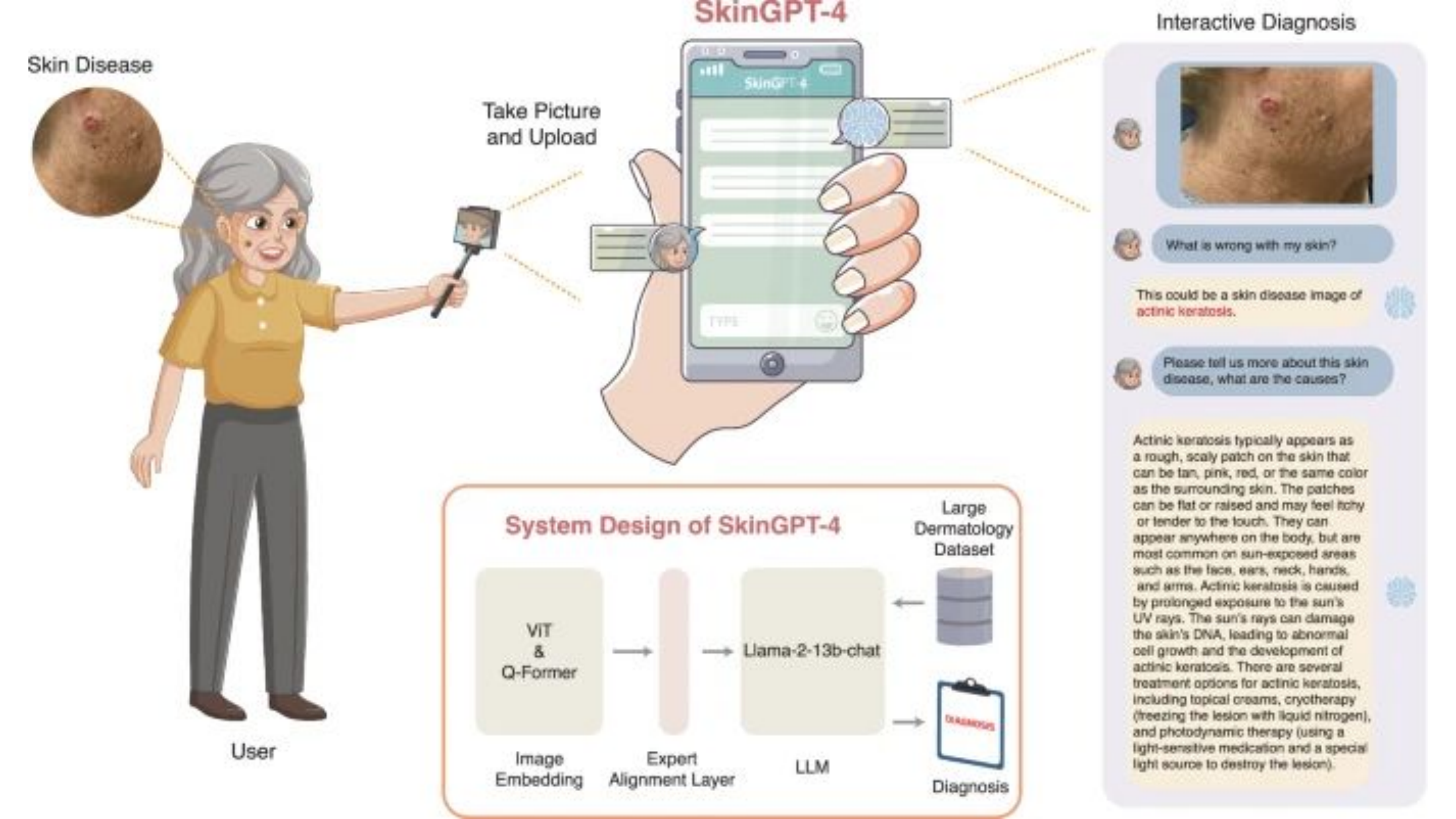
Nature Communications **15**, Article number: 5649 (2024) | Cite this article

Abstract

Large language models (LLMs) are seen to have tremendous potential in advancing medical diagnosis recently, particularly in dermatological diagnosis, which is a very important task as skin and subcutaneous diseases rank high among the leading contributors to the global burden of nonfatal diseases. Here we present SkinGPT-4, which is an interactive dermatology diagnostic system based on multimodal large language models. We have aligned a pre-trained vision transformer with an LLM named Llama-2-13b-chat by collecting an extensive collection of skin disease images (comprising 52,929 publicly available and proprietary images) along with clinical concepts and doctors' notes, and designing a two-step training strategy. We have quantitatively evaluated SkinGPT-4 on 150 real-life cases with board-certified dermatologists. With SkinGPT-4, users could upload their own skin photos for diagnosis, and the system could autonomously evaluate the images, identify the characteristics and categories of the skin conditions, perform in-depth analysis, and provide interactive treatment recommendations.

Pre-trained Multimodal Large Language Model Enhances Dermatological Diagnosis Using SkinGPT-4

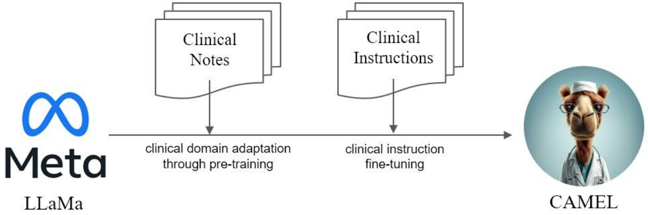
- Main Task:
  - Develop a system that improves dermatological diagnosis using SkinGPT-4, leveraging multiple modalities (images and text).
- Approaches:
  - SkinGPT-4 integrates a pre-trained Vision Transformer with Llama-2-13b-chat.
  - Trained on 52,929 skin disease images with clinical inputs.
- Results:
  - Evaluated on 150 real-life cases; consistently accurate in diagnosis.
  - Offers immediate, interactive, and autonomous image analysis and treatment recommendations.
  - Provides faster response times compared to traditional dermatology consultations.
  - High user and expert satisfaction regarding diagnosis accuracy and user privacy.



SkinGPT-4
Skin Disease
Take Picture
and Upload
SkinGPT-4
TYPE
User
Interactive Diagnosis
What is wrong with my skin?
This could be a skin disease image of actinic keratosis.
Please tell us more about this skin disease, what are the causes?
Actinic keratosis typically appears as a rough, scaly patch on the skin that can be tan, pink, red, or the same color as the surrounding skin. The patches can be flat or raised and may feel itchy or tender to the touch. They can appear anywhere on the body, but are most common on sun-exposed areas such as the face, ears, neck, hands, and arms. Actinic keratosis is caused by prolonged exposure to the sun's UV rays. The sun's rays can damage the skin's DNA, leading to abnormal cell growth and the development of actinic keratosis. There are several treatment options for actinic keratosis, including topical creams, cryotherapy (freezing the lesion with liquid nitrogen), and photodynamic therapy (using a light-sensitive medication and a special light source to destroy the lesion).
System Design of SkinGPT-4
ViT
&
Q-Former
Image
Embedding
Expert
Alignment Layer
Llama-2-13b-chat
LLM
Large
Dermatology
Dataset
DIAGNOSIS
Diagnosis

Clinically Adapted Model Enhanced from LLaMA (CAMEL)

- Main Task:
  - Develop a privacy-preserving clinical language model to support healthcare decision-making.
- Approach:
  - CAMEL is based on LLaMA, pre-trained on MIMIC-III/IV notes (3.4B tokens) (Medical Information Mart for Intensive Care).
  - Fine Tuned on 100,000 clinical instructions, focused on 13 NLP clinical tasks.
- Preliminary Results:
  - CAMEL achieved 96% of GPT-3.5's performance, as assessed by GPT-4.
  - Demonstrated superior performance compared to Alpaca (LLaMA based model from Stanford) in 80% of the tasks.
  - Released on PhysioNet for credentialized access; replicable using in-house clinical notes.



https://starmpcc.github.io/CAMEL/

https://physionet.org/

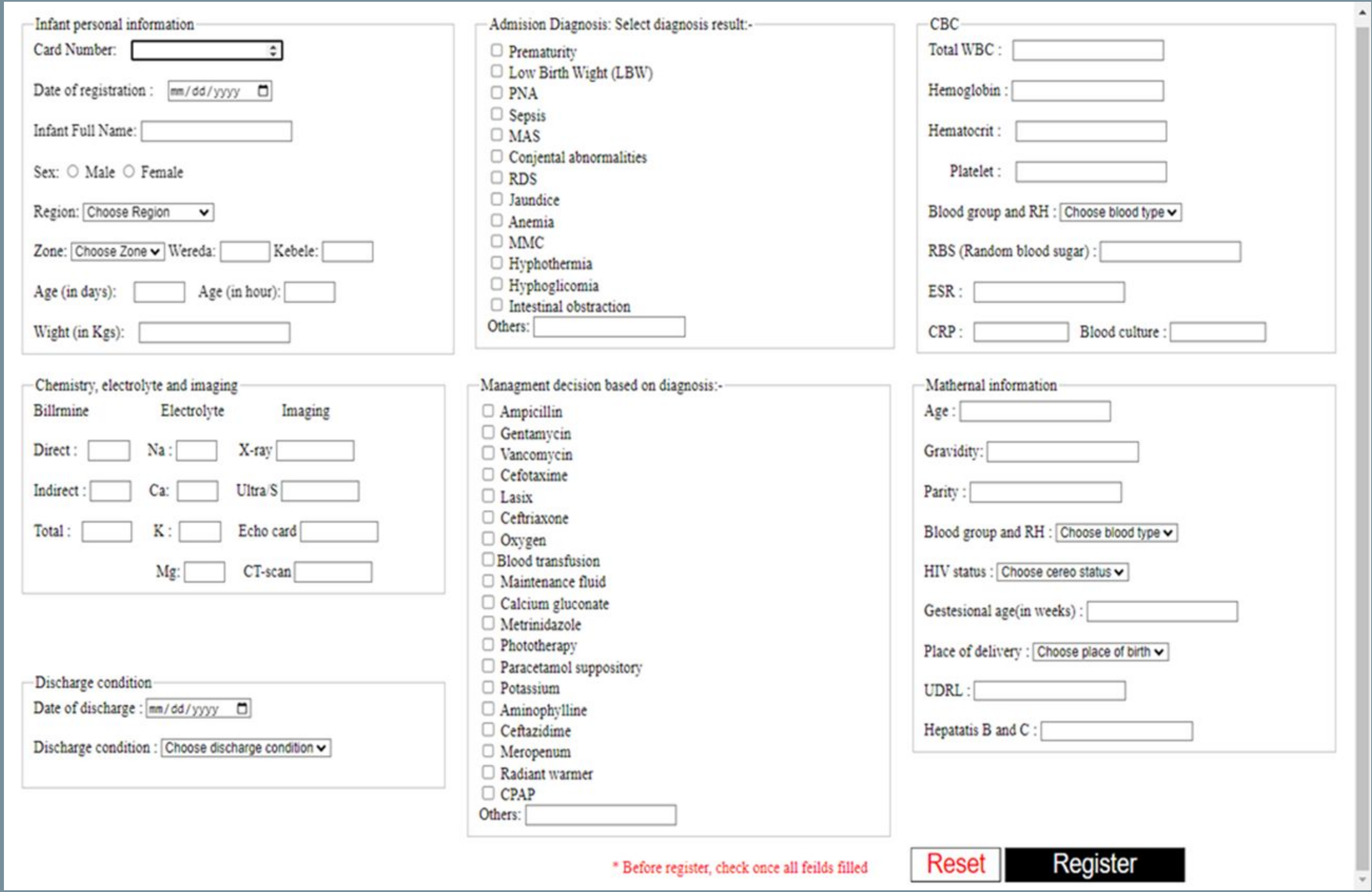
Camel Demo

https://starmpcc-camel-demo-demo-i7ajms.streamlit.app/

Some of Our Works and Projects

Data Science and AI Approaches for Neonatal Health Data

- Motivations:
  - How time serious data will be used to forecast neonatal mortality?
  - What are the determinant factors of neonatal mortality in Ethiopia?
  - How to implement machine learning algorithms to mortality and APGAR score prediction?
- Data collection:
  - From Sep 2022 to Jun 2023.
  - 3026 records with 44 features



Infant personal information

Card Number:

Date of registration : mm/dd/yyyy

Infant Full Name:

Sex: ○ Male ○ Female

Region: Choose Region

Zone: Choose Zone Wereda: Kebele:

Age (in days): Age (in hour):

Wight (in Kgs):

Admision Diagnosis: Select diagnosis result:-

- ☐ Prematurity
- ☐ Low Birth Wight (LBW)
- ☐ PNA
- ☐ Sepsis
- ☐ MAS
- ☐ Conjental abnormalities
- ☐ RDS
- ☐ Jaundice
- ☐ Anemia
- ☐ MMC
- ☐ Hyphothermia
- ☐ Hyphoglicomia
- ☐ Intestinal obstraction

Others:

CBC

Total WBC :

Hemoglobin :

Hematocrit :

Platelet :

Blood group and RH : Choose blood type

RBS (Random blood sugar) :

ESR :

CRP : Blood culture :

Chemistry, electrolyte and imaging

Billrmine	Electrolyte	Imaging
Direct :	Na :	X-ray
Indirect :	Ca:	Ultra/S
Total :	K :	Echo card
	Mg:	CT-scan

Managment decision based on diagnosis:-

- ☐ Ampicillin
- ☐ Gentamycin
- ☐ Vancomycin
- ☐ Cefotaxime
- ☐ Lasix
- ☐ Ceftriaxone
- ☐ Oxygen
- ☐ Blood transfusion
- ☐ Maintenance fluid
- ☐ Calcium gluconate
- ☐ Metrinidazole
- ☐ Phototherapy
- ☐ Paracetamol suppository
- ☐ Potassium
- ☐ Aminophylline
- ☐ Ceftazidime
- ☐ Meropenum
- ☐ Radiant warmer
- ☐ CPAP

Others:

Mathernal information

Age :

Gravidity:

Parity :

Blood group and RH : Choose blood type

HIV status : Choose cereo status

Gestesional age(in weeks) :

Place of delivery : Choose place of birth

UDRL :

Hepatatis B and C :

Discharge condition

Date of discharge : mm/dd/yyyy

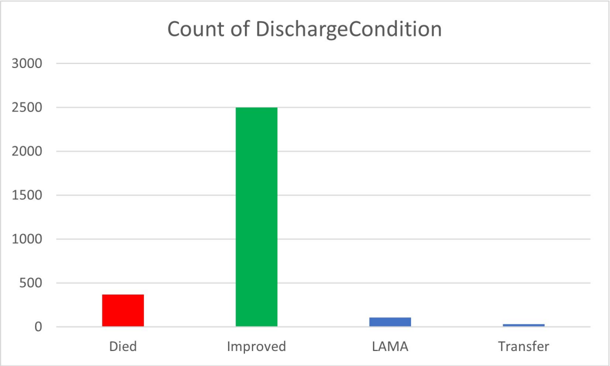
Discharge condition : Choose discharge condition

* Before register, check once all feilds filled

Reset Register

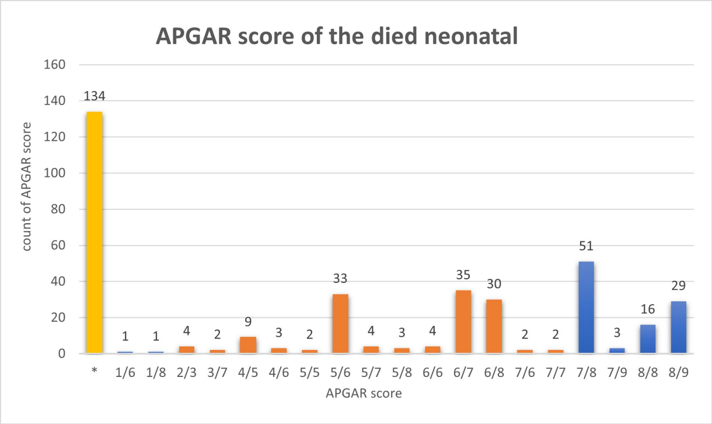
Data analysis

- Discharge condition: from a total of 3009 records, 2501 are improved, 368 died, 108 LAMA, and transfer 32.
- **Place of delivery for dead neonatal:** 174 from clinic, 38 home, 158 hospital (5.78%, 1.26%, 5.25%)



APGAR: quick test on a baby at 1’ and 5’ of birth.

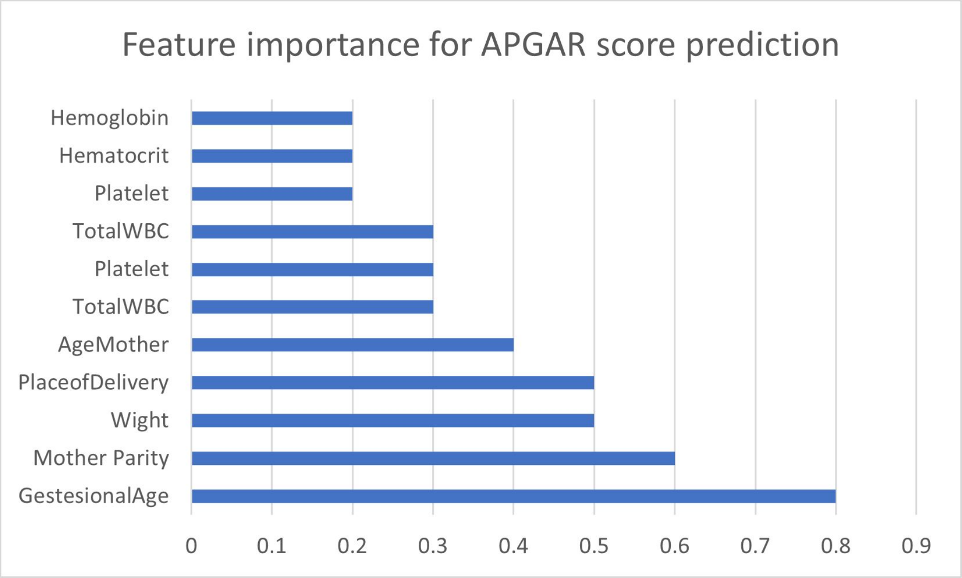
- Ranges from 0 to 10,
- 7-10 normal
- 4-6 needs proper supervision
- <=3 is never good.



- In 2015 E.C., for DCSH, the neonatal mortality rate is around 12%.
- Predict neonatal APGAR score:

Classifier	Precision	Recall	F1 Score	Accuracy
SVM	94.3	92.4	93.4	95.7
RF	**96.7**	**97.0**	**96.9**	**96.69**

- Top risk factors for the cause of neonatal death



Highly risky diseases

- Sepsis,
- LBW,
- Hypothermia,
- prematurity,
- RDS,
- PNA

Adaption and Evaluation of Generative Large Language Models for German Medical Information Extraction

- **Research Questions:**
  - Can LLMs with **7 billion parameters** compete with smaller, fine-tuned models (**SLMs**) for **German medical information extraction**?
  - How can an NLP pipeline-based application enhance **clinical document analysis** for premedication reports?
- **Methodology:**
  - Evaluated 7 billion parameter LLMs on German clinical datasets.
  - Developed the **MEDICA app** prototype for supporting physicians in premedication processes.
- **Solutions:**
  - Introduced **Instruction Tuning** using QLoRA to improve extraction performance.
  - Utilized a **user study to validate** MEDICA's ability to streamline premedication reports.

1. Dr. Seid Muhie Yimam (First reviewer + Thesis Supervisor)
2. Prof. Dr. Frank Ückert (Second reviewer)

Spiegel, S. (2024). Adaption and Evaluation of Generative Large Language Models for German Medical Information Extraction [Master's thesis, Universität Hamburg].

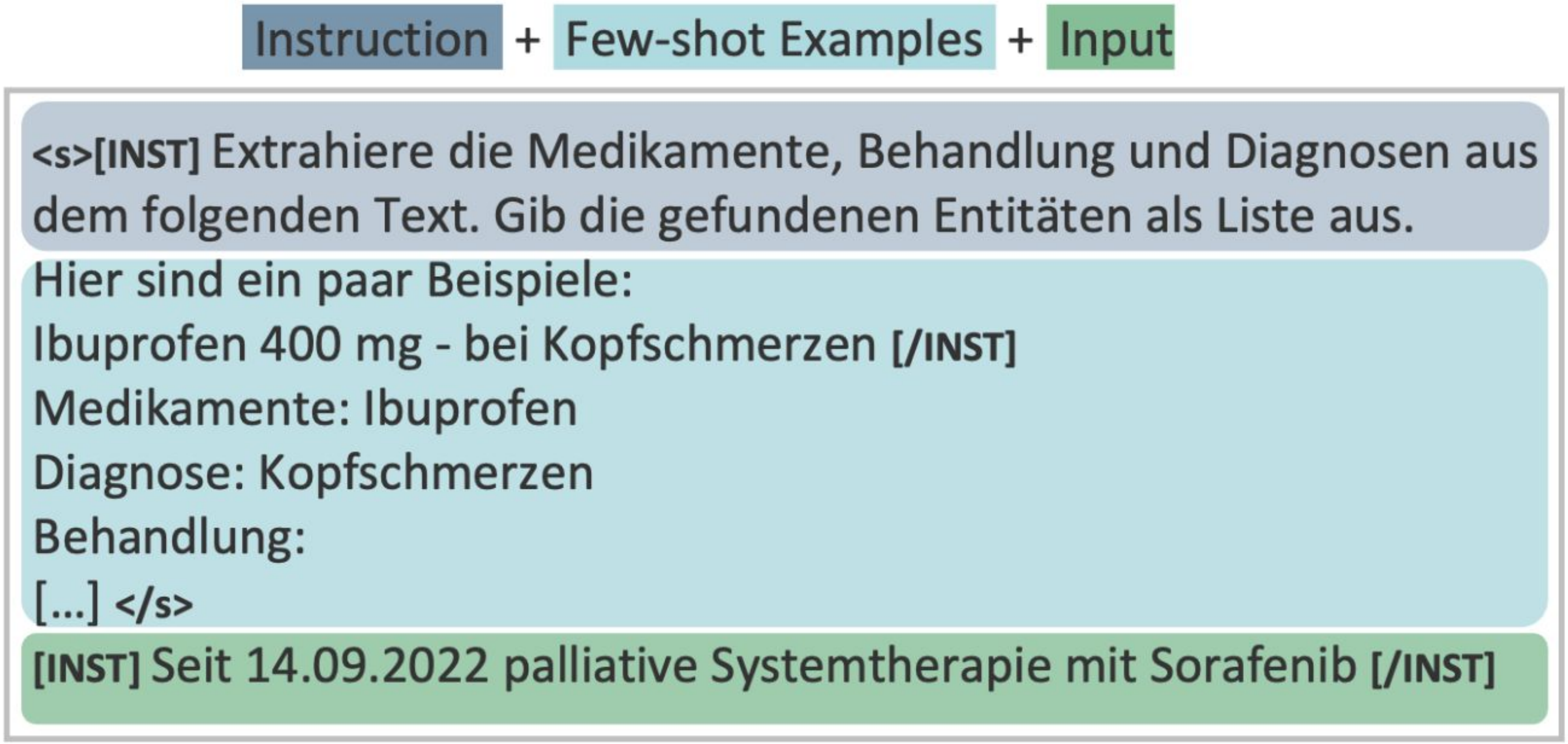


Figure 4.1.: Few-shot approach to extract named entities from input

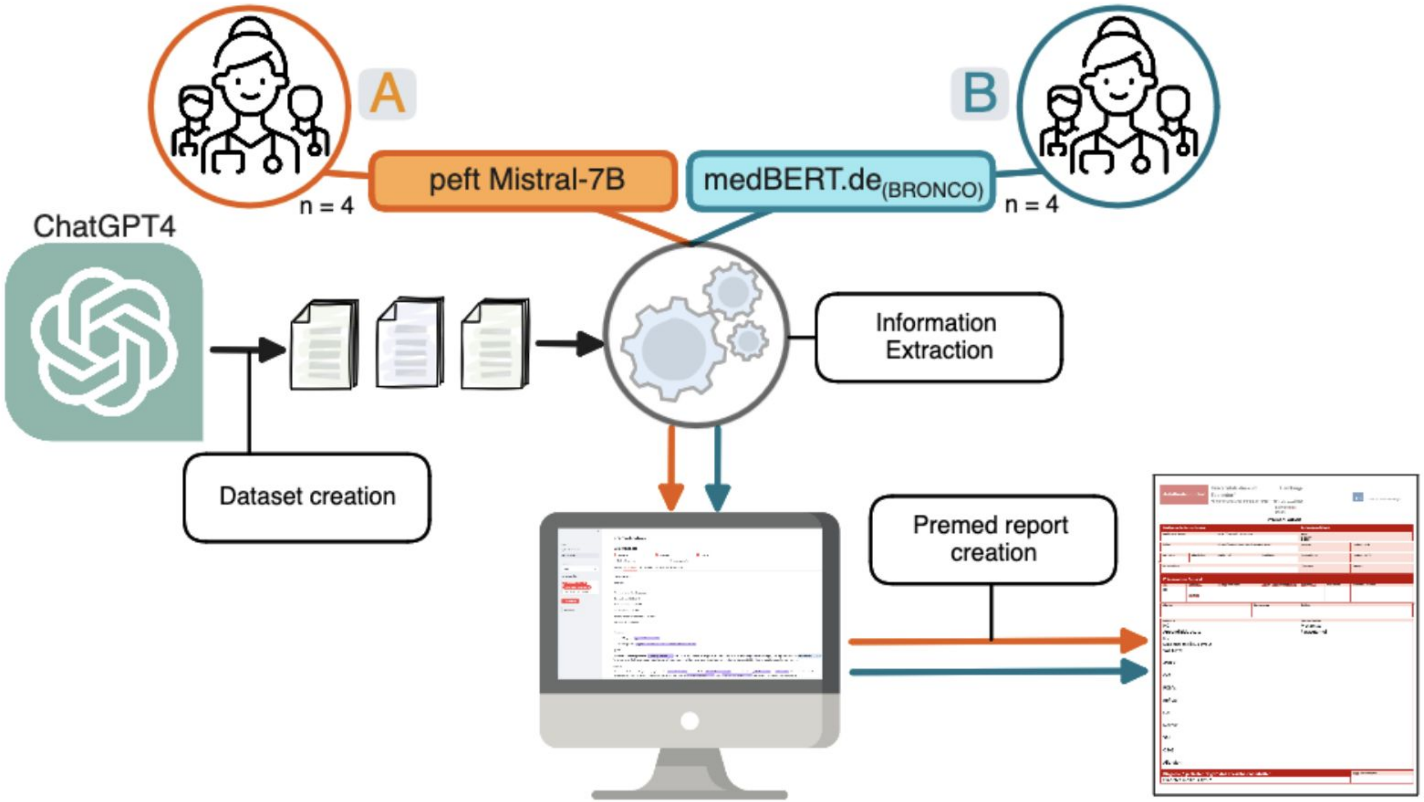
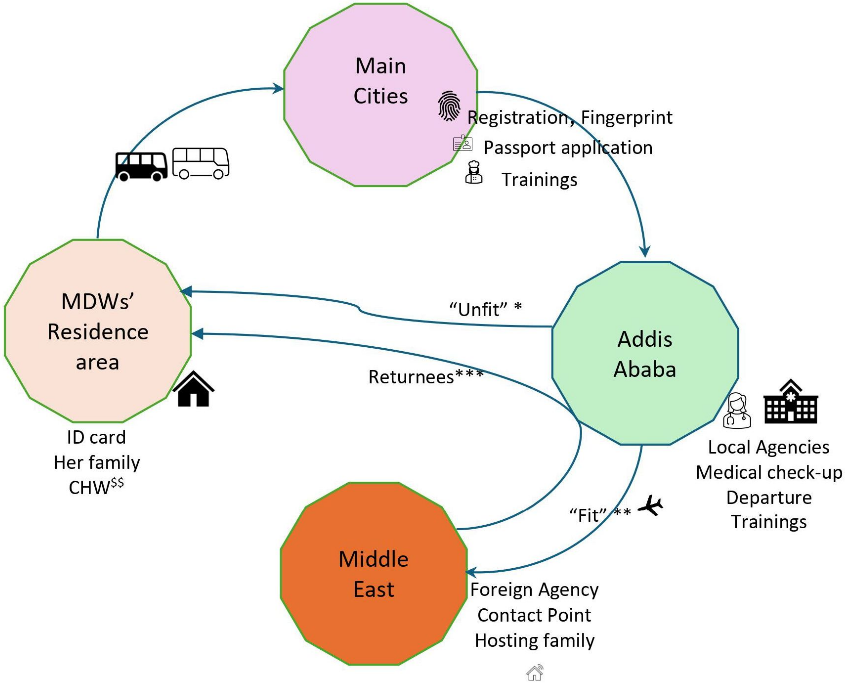


Figure 4.5.: Use study design to evaluate the developed MEDICA application in the context of the premedication report creation.

INNOVETH: Health Intervention for Ethiopian MDWs

- **Objective**:
  - To improve health outcomes for Ethiopi Migrant Domestic Workers (MDWs) in t Middle East by leveraging big data and technology.
- **Key Strategies:**
  - **Epidemiological** and **data science** methods to identify health challenges.
  - Risk stratification models for mental, sexual, and reproductive health.
  - Transdisciplinary intervention package (counseling, peer-support, self-help groups, mobile apps).
  - Evaluation through randomized control trials.

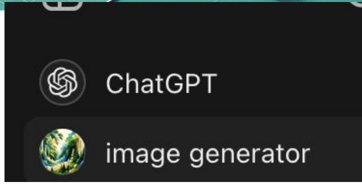


INNOVETH Project Implementation

- Phase I: Assessment and Mapping
  - Identify health issues through systematic reviews and qualitative methods.
  - Utilize migration data for health problem identification and risk assessment.
- Phase II: Intervention and Evaluation
  - Develop tailored interventions based on Phase I findings.
  - Implement with randomized control trials for impact and cost-effectiveness evaluation.
- Expected Outcomes:
  - Improved health and well-being of MDWs.
  - Strengthened resilience and social networks.
  - Enhanced policy impact and inclusivity in health programs.

Questions and Discussion

- Low Resource Setup
  - How can LLMs work in low-resource healthcare settings?
  - What are efficient strategies for using LLMs with limited resources?
- Hospital System Automation
  - What barriers exist to digitalizing hospitals?
  - How can LLMs help automate hospital tasks?
- Career Impact
  - Will LLMs replace jobs in healthcare?
  - What skills are needed to work alongside AI in healthcare?
- International Collaboration
  - How can global cooperation enhance healthcare with AI?



Contact us

Dr. M Sanni Ali, London School of Hygiene and Tropical Medicine,UK
Sanni.Ali@lshtm.ac.uk

Dr. Seid Muhie Yimam, University of Hamburg, HCDS, DE
seid.muhie.yimam@uni-hamburg.de