Knowledge Graphs and LLMs in Rare Diseases

14th April

LifeArc

DRIVE-Health
EPSRC Centre for Doctoral Training in Data-Driven Health

UKRI Engineering and Physical Sciences Research Council

King's College London

[Disclaimer]

This talk represents my views and not the views of my funders.

[Overview]

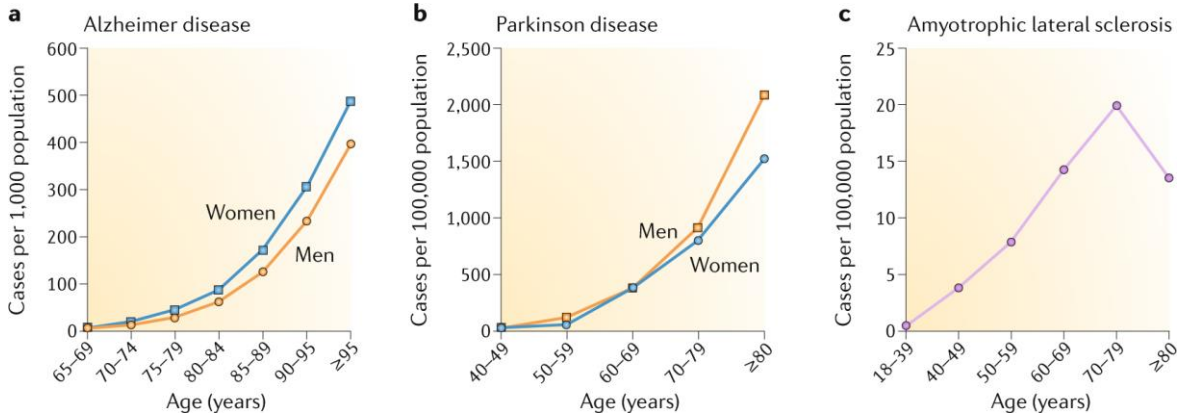
1. The rare disease problem
2. Why LLMs alone fall short
3. What KGs bring
4. Some research looking into LLMs, KGs and both in rare disease research
5. Open challenges and what’s next?

[Overview]

1. The rare disease problem
2. Why LLMs alone fall short
3. What KGs bring
4. Some research looking into LLMs, KGs and both in rare disease research
5. Open challenges and what's next?

[What's the rare disease problem?]

- A LOT of rare diseases with huge genetic component
- Rare neurodegenerative diseases: Parkinson's, dementias, MND, rare metabolic disorders
- Rare diseases are hard to diagnose, and treat
- The knowledge exists — but fragmented, not connected, not queryable at the point of care



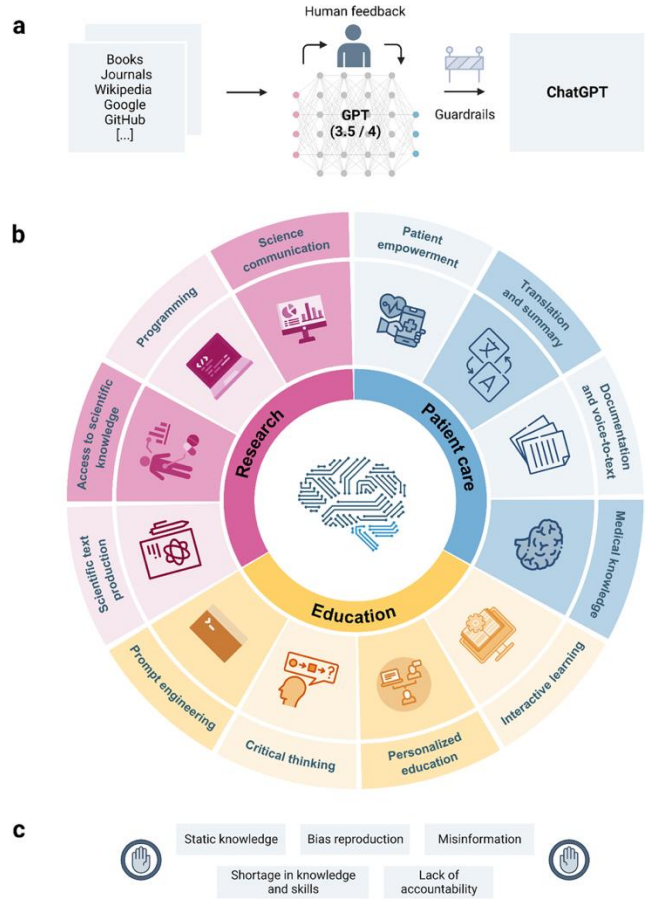
Hou, Y., Dan, X., Babbar, M. *et al.* Ageing as a risk factor for neurodegenerative disease. *Nat Rev Neurol* **15**, 565–581 (2019). https://doi.org/10.1038/s41582-019-0244-7

[Overview]

1. The rare disease problem
2. Why LLMs alone fall short
3. What KGs bring
4. Some research looking into LLMs, KGs and both in rare disease research
5. Open challenges and what's next?

[Why LLMs alone fall short]

1. LLMs trained on internet-scale text: skewed heavily toward common diseases
2. Rare neurodegenerative diseases = data desert in training corpora
3. Multiple failure areas
4. Alternative view: rare disease diagnosis is a graph traversal problem, not a language problem



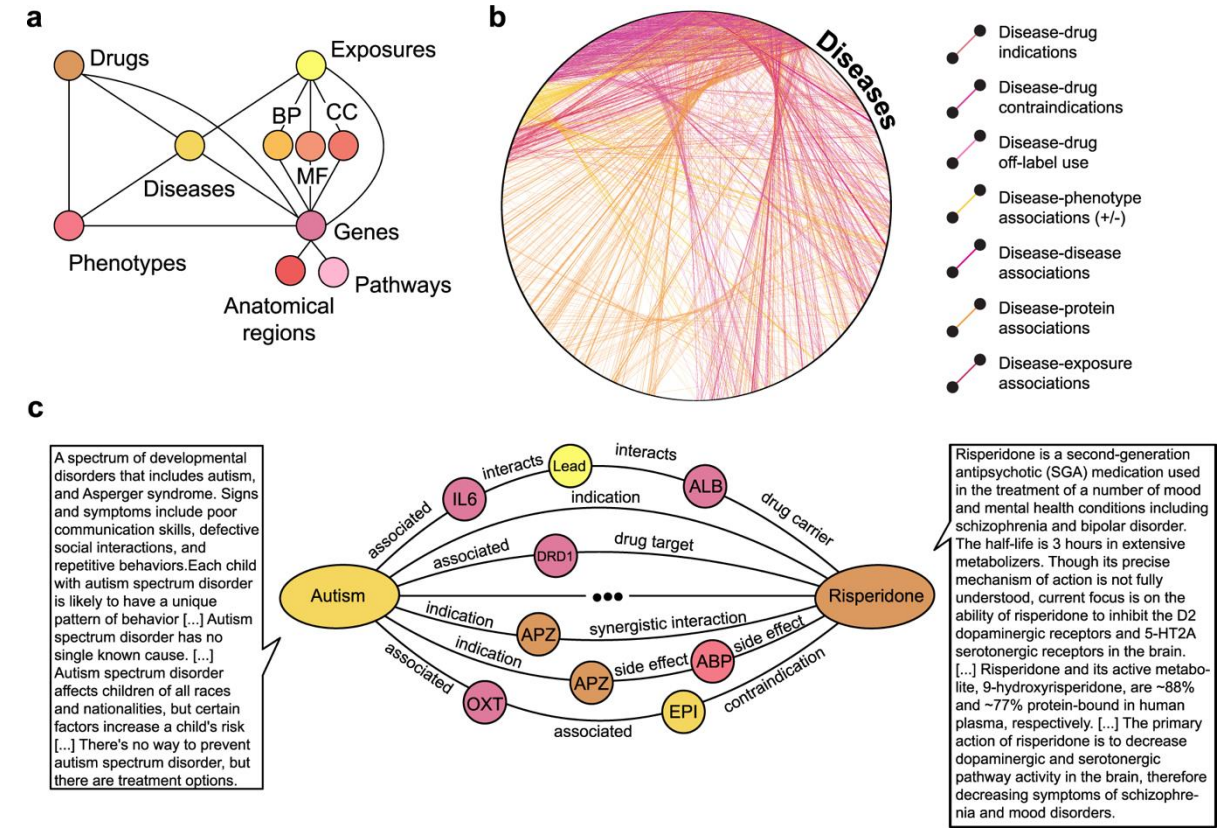
Clusmann, J., Kolbinger, F.R., Muti, H.S. *et al.* The future landscape of large language models in medicine. *Commun Med* **3**, 141 (2023). https://doi.org/10.1038/s43856-023-00370-1

[Overview]

1. The rare disease problem
2. Why LLMs alone fall short
3. What KGs bring
4. Some research looking into LLMs, KGs and both in rare disease research
5. Open challenges and what’s next?

[What do KGs bring?]

- Key biomedical ontologies
- Phenotype → gene → disease triples
- Multi-hop reasoning
- Explainable paths
- Limitation: incomplete, slow to update, can't handle unstructured clinical text — could use LLMs



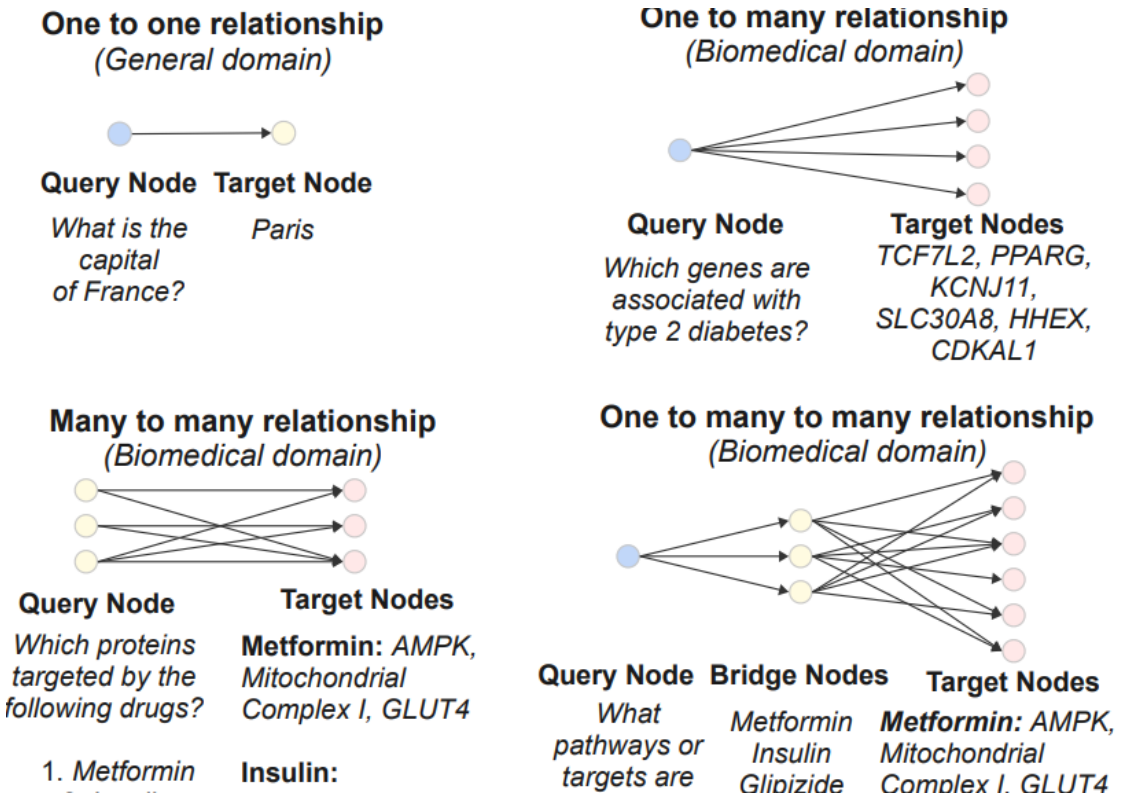
Chandak, P., Huang, K. & Zitnik, M. Building a knowledge graph to enable precision medicine. *Sci Data* **10**, 67 (2023). https://doi.org/10.1038/s41597-023-01960-3

[Overview]

1. The rare disease problem
2. Why LLMs alone fall short
3. What KGs bring
4. Some research looking into LLMs, KGs and both in rare disease research
5. Open challenges and what’s next?

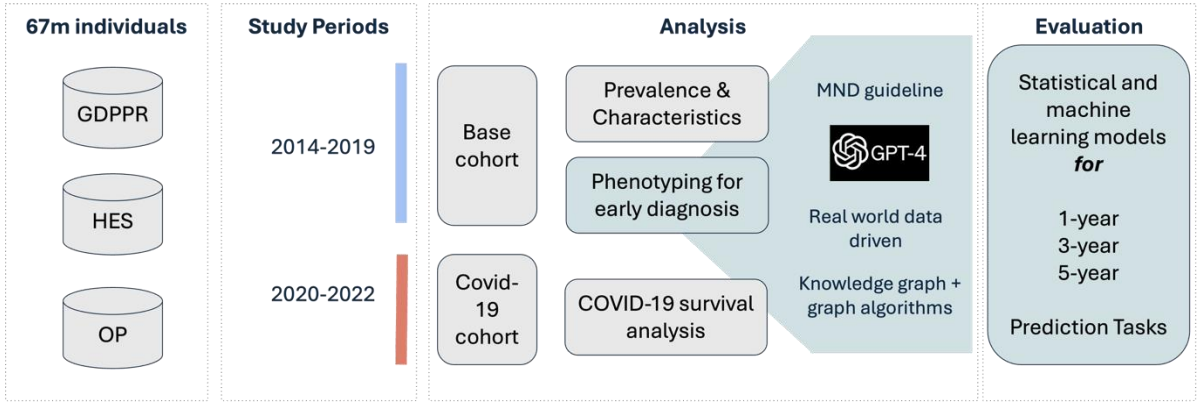
[BioHopR: benchmarking multi-hop reasoning in biomedicine]

- Built on PrimeKG: 1-hop and 2-hop QA tasks across drug → disease → protein chains
- Best model (O3-mini): 37.93% on 1-hop, 14.57% on 2-hop — all models fail sharply at two steps
- LLMs cannot yet reliably traverse the multi-hop paths that rare disease diagnosis requires — KGs are necessary scaffolding, not optional context



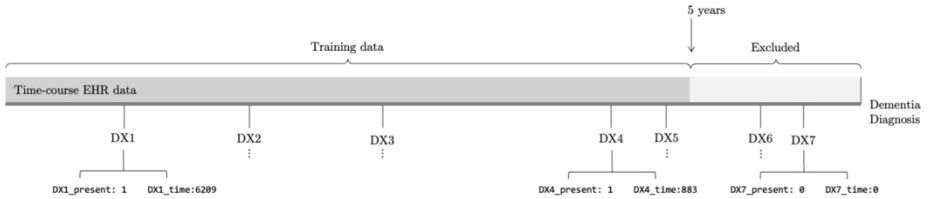
[Early diagnosis of MND across 67 million people in England]

- 12,240 MND cases from linked NHS England records
- KG (PrimeKG + PageRank) identified predictive areas: digestive, respiratory, mental disorders — all absent from guidelines
- Ensemble of KG + guideline + real-world data achieved 28% F1 improvement — structured knowledge and data-driven approaches are stronger together



[Predicting dementia risk from longitudinal EHR using graph neural networks]

- Each patient represented as a graph: diagnoses as nodes, time since diagnosis as edge weights — capturing not just what but when
- GNN outperforms BERT and all baselines — AUCROC 0.838 after risk stratification, top predictors flagged years before diagnosis
- Graph-structured reasoning over longitudinal EHR captures what flat models miss — a natural bridge between KGs and real-world clinical prediction



[Overview]

1. The rare disease problem
2. Why LLMs alone fall short
3. What KGs bring
4. Some research looking into LLMs, KGs and both in rare disease research
5. Open challenges and what’s next?

[Open challenges and what's next?]

1. Reasoning gap: LLMs still fail at multi-hop traversal — KG-grounding is needed to guide reasoning, not just provide context

2. Temporal knowledge: static KGs go stale, LLMs could help update

3. Clinical translation: embedding KG-grounded LLMs into real-world decision support for rare neurodegeneration

[Acknowledgements / Thank you]

PHIDL
Precision Health
Informatics Data Lab

UK MND
Research Institute