MEDINFO 2023 — The Future Is Accessible
J. Bichel-Findlay et al. (Eds.)

doi:10.3233/SHTI231207

A Personal Healthcare Knowledge Graph Framework for Diagnosis of Pelvic Masses Diseases

Ran Xin[a,b], Jia Chen[b], Feifei Bao[b], Yong Shang[b], Xu Han[b] and Jingsong Li[b,c,1]
[a] *Polytechnic Institute, Zhejiang University, Hangzhou, China*
[b] *Research Center for Healthcare Data Science, Zhejiang Laboratory, Hangzhou, China*
[c] *Engineering Research Center of EMR and Intelligent Expert Systems, Ministry of Education, College of Biomedical Engineering and Instrument Science, Zhejiang University, Hangzhou, China*

Abstract. A Personal Health Knowledge Graph (PHKG) facilitates the efficient integration of potential diagnostic clues from patients' electronic health records with medical knowledge, establishing diagnostic reasoning paths and ensuring accurate, individually interpretable results in the diagnosis of pelvic masses.

Keywords. Personal health knowledge graph, interpretable reasoning, ontology, pelvic masses, quadruplet.

1. Introduction

Pelvic masses, common in gynecology, pose a diagnostic challenge due to their association with diverse diseases and the need for synthesizing examination results. The Personal Health Knowledge Graph (PHKG) combines personal health data and medical knowledge into a structured graph, facilitating information integration and visualizing individual reasoning, offering potential for addressing pelvic mass diagnostics[1, 2].

2. Methods

This study utilizes data from 22,303 patients with pelvic masses admitted to Zhejiang Women's Hospital from January 2014 to December 2020. Electronic health records were processed through information extraction and semantic transformation into static triplets, utilizing natural language processing for knowledge fusion and linking with expert-curated medical knowledge. The accumulated information forms a PHKG, where a Reasoning Path (RP) is depicted as a potential multi-hop path from person (P) to disease (D) (Figure 1).

The risk of pelvic mass-related diseases, denoted as RS, is calculated via reasoning path (RP) score, based on the formula $RS(P,D) = IS(P,D) * \sum_{t \in \mathcal{T}} PS(t)$; $PS(t) = \prod_{r \epsilon t} w(r)$. Here, w(r) signifies the triplet score within the RP, and IS pertains to the

[1] Corresponding Author: Jingsong Li (e-mail: ljs@zju.edu.cn).

completeness of potential diagnostic indicators during medical consultation. Scoring each PHKG triplet involves assigning credibility based on information source reliability, scoring rule-generated triplets (e.g., lab tests) by abnormality, and calculating entity-linked triplet similarity scores using a pre-trained language model. Expert suggestions supplement annotations for relationships from clinical consensus sources, and TF-IDF referencing identifies potential node category associations. These results form edge weights in the reasoning path quadruplets. The framework's diagnostic reasoning performance was assessed for multiple and single disease types using NDCG@k, MRR@k, and F1-score metrics.

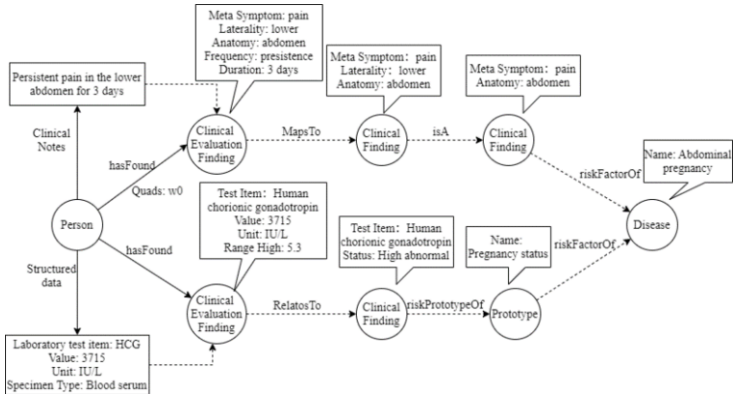

Figure 1: An example of PHKG construction procedure

3. Results

Our study included 22,303 patients diagnosed with 41 diseases, with benign ovarian tumors, tubal pregnancy, and ovarian endometriosis cysts being the top three. We used a 7:3 training-to-test set ratio for model construction and evaluation. Our constructed PHKG of pelvic masses contained over 30131 entities and nearly 3327677 million triplets. Performance evaluation on the test set showed NDCG@3 and MRR@3 scores of 0.836 and 0.804 across 41 diseases, and F1-scores of 0.855, 0.882, and 0.889 for the top three diseases, respectively.

4. Conclusions

This study shows that with quantitative and interpretable inference path presented for each patient, the proposed PHKG framework demonstrates a good potential in diagnosing pelvic masses.

References

[1] S. Shirai, O. Seneviratne, and D.L. McGuinness, Applying Personal Knowledge Graphs to Health, *ArXiv210407587 Cs*. (2021). http://arxiv.org/abs/2104.07587 (accessed October 12, 2021).

[2] N. Rastogi, and M.J. Zaki, Personal Health Knowledge Graphs for Patients, *ArXiv200400071 Cs*. (2020). http://arxiv.org/abs/2004.00071 (accessed October 25, 2021).