The **Healthcare Knowledge Graph powered by PubMed** literature can help Solara Health evaluate how effective a given **treatment protocol** or **formulary** is for members by synthesizing and reasoning over evidence in the literature. Here's how we can approach it conceptually and practically:

**Key Idea:**

We build a Knowledge Graph that connects concepts such as:

* **Diseases / Conditions**
* **Medications / Therapies**
* **Clinical Protocols / Guidelines**
* **Outcomes / Endpoints**
* **Patient Populations**
* **Adverse Events**
* **Efficacy Measures**  
  based on co-occurrence, causal relationships, and statistical associations mined from **PubMed abstracts and full text**.

**Use Case Example**

Suppose Solara Health wants to evaluate whether **Protocol A** (e.g., prescribing a specific statin + behavioral counseling) is effective for **diabetic members aged 45–65**.

**Graph Model Sketch**

(:Protocol)-[:RECOMMENDS]->(:Drug)

(:Protocol)-[:TARGETS]->(:Condition)

(:Drug)-[:MENTIONED\_IN]->(:PubMedArticle)

(:PubMedArticle)-[:SHOWS\_EFFECT\_ON]->(:Outcome)

(:Outcome)-[:HAS\_MEASURE]->(:EfficacyMetric)

(:Population)-[:DESCRIBED\_IN]->(:PubMedArticle)

**Sample Questions This Graph Can Answer**

| **Question** | **Cypher-style (simplified)** |
| --- | --- |
| What outcomes are associated with Protocol A? | MATCH (p:Protocol {name:"Protocol A"})-[:RECOMMENDS]->(d:Drug)<-[:MENTIONED\_IN]-(a:Article)-[:SHOWS\_EFFECT\_ON]->(o:Outcome) RETURN o |
| Is there evidence that Protocol A improves A1C in diabetic patients? | Add a filter on (o:Outcome {type: "A1C reduction"}) and Population nodes |
| What are common adverse events linked to formulary drugs in Protocol A? | Expand with (:Drug)-[:CAUSES]->(:AdverseEvent) |
| Is there evidence supporting this protocol in populations similar to ours? | Use embeddings or similarity on Population nodes and compare |
| How does this compare to Protocol B? | Do a path-based comparison of Protocol A vs. Protocol B effectiveness paths |

**Techniques Involved**

1. **NER + Entity Linking** to identify drugs, diseases, outcomes, populations from PubMed.
2. **Relation Extraction (RE)** to find causal links like “Drug X reduced blood pressure.”
3. **Graph Construction** using Neo4j or similar, connecting entities and relationships.
4. **Query / Path Traversal** to reason about outcomes, efficacy, and side effects.
5. **Embedding Similarity (Graph ML)** to find similar populations or treatments.
6. **Evidence Scoring** based on number of articles, confidence, journal impact, etc.

**Benefits for Solara Health**

* **Evidence-Based Protocol Optimization**: Justify changes in treatment guidelines.
* **Population Matching**: Ensure evidence aligns with *our members* (e.g., age, comorbidities).
* **Formulary Design**: Choose drugs supported by strongest real-world and clinical evidence.
* **Risk Management**: Understand adverse effect risks at population scale.