Rare diseases pose particular challenges to patients, families, caregivers, clinicians and researchers. Due to the scarce availability of information and their disintegration, in recent years we are witnessing a strong growth of patient communities on social platforms such as Facebook. Within them, users share experiences and seek answers to questions aimed at improving their health conditions (e.g., “which are the most effective medical treatments from patients’ viewpoint?”). Though the unlabeled data generated in this context are of high value, the currently existing ontologies and resources tend to ignore them. Moreover, the explanation of a phenomenon from text documents is still a difficult and poorly addressed problem.

**Aims**

- A novel methodology of descriptive text mining, capable of offering accurate explanations in unsupervised settings, of discovering relationships among concepts and quantifying them according to their statistical significance (e.g., “citrus fruit” ↔ “acid reflux”: 87%).
- Introduction of a new knowledge graph learning approach.
- Effectiveness demonstration through a case study focused on Esophageal Achalasia (ORPHA: 930), extending ORDO and introducing a patient-centered knowledge graph into the world of Linked Open Data.

**Patient Knowledge Extraction**

- Documents Preprocessing
- Language Model Application (terms and documents)
- IR methods for correlations detection
- Documents Classification
- Terms Weighting
- Phenomenon explanation based on hypothesis tests

**Patient Knowledge Representation**

- Positive and negative explanations for Achalasia treatments (Table 1).

<table>
<thead>
<tr>
<th>Heller-Dor</th>
<th>POEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>equipe, dr mario, costantini, salvador, padua, antireflux, plastic problems, drink, eat</td>
<td></td>
</tr>
<tr>
<td>reflux, problems, liquid, pain, inflammation, antacid</td>
<td></td>
</tr>
</tbody>
</table>

**Experiments and Results**

- **Explanations effectiveness.** Successful recognition of the positive and negative aspects of the two main intervention techniques for Esophageal Achalasia (Heller-Dor and POEM) through local analyzes and p-value < 0.01 (Table 1).

<table>
<thead>
<tr>
<th>Confusion Matrix</th>
<th>Statistical Indices</th>
</tr>
</thead>
<tbody>
<tr>
<td>P-value Threshold</td>
<td>TP</td>
</tr>
<tr>
<td>0.7</td>
<td>101</td>
</tr>
<tr>
<td>0.8</td>
<td>96</td>
</tr>
<tr>
<td>0.9</td>
<td>72</td>
</tr>
</tbody>
</table>

**Conclusions and Future Works**

- Combined use of Text Mining and Semantic Web techniques, with a modular, automatic and domain-independent KE e KR methodology.
- Identification of statistical evidences and scientific medical correlations directly from patients’ posts accumulated over time, avoiding their manual reading.
- The methodology could be applied to different communities of patients for Esophageal Achalasia (comparing the results) and to other diseases.

**References**

This poster abstract is a reworking of two contributions.