

## Analysis of Causal Relationships in Integrated Ontologies of Diseases, Phenotypes, and Radiological Diagnosis

Charles E. Kahn, Jr.

Department of Radiology and Institute for Biomedical Informatics,  
University of Pennsylvania, Philadelphia, Pennsylvania, USA

### Abstract

Biomedical ontologies encode knowledge in a form that makes it computable. The current study used the integration of three large biomedical ontologies—the Disease Ontology (DO), Human Phenotype Ontology (HPO), and Radiology Gamuts Ontology (RGO)—to explore inferred causal relationships between high-level DO and HPO concepts. The principal DO categories were defined as the 7 direct subclasses of the top-level Disease class, excluding Disease of anatomical entity, plus the 12 direct subclasses of the latter term. The principal HPO categories were defined as the 25 direct subclasses of HPO's Phenotypic abnormality class. All causal relationships were tallied between members of the DO and HPO principal categories through their causal relationships in RGO. The analysis provides an understanding of the hierarchical organization of RGO terms, and offers insights into new relationships between DO and HPO classes.

### Keywords:

Artificial Intelligence; Biological Ontologies; Knowledge Bases.

### Introduction

An ontology is a knowledge-representation construct that describes a domain of interest with a set of logical statements, or axioms. Those axioms relate one concept within the ontology to another. An ontology classifies the entities within a domain; each entity is said to make up a term, or “class,” of the ontology [13]. The ontology approach has the advantage that it expresses knowledge in a form that humans can read and automated systems can process [1]. Thus, an ontology can allow both humans and computers to describe and reason about the concepts in a domain. Ontologies help promote clarity and can enable disparate medical systems to work together [14].

Medical imaging plays a key role in diagnosis and treatment, and a patient's imaging findings compose part of that person's phenotype. Thus, understanding a patient's imaging phenotype can enable precision-medicine approaches that enable targeted therapies based on an individual's genomic, epigenomic, and gene-expression patterns. The Radiology Gamuts Ontology (RGO) encodes relationships between imaging findings and diagnoses to offer a formal representation of knowledge of diagnostic radiology [2].

This study analyzed one aspect of new knowledge generated through integration of ontologies of diseases, phenotypes, and radiological diagnosis. We sought to understand potential linkages between high-level concepts in ontologies of diseases and

phenotypes as related through their relationships in an ontology of imaging findings.

### Methods

#### Ontologies

Three biomedical ontologies—the Disease Ontology, Human Phenotype Ontology, and Radiology Gamuts Ontology—formed the basis for the current analysis; no patients or protected health information was involved.

#### Disease Ontology

The Disease Ontology (DO, version 2020-11-11) offers a hierarchically organized vocabulary of 17 375 diseases that afflict humans [8; 12]. DO incorporates cross-mapped concepts from widely used biomedical terminologies, such as International Classification of Diseases (ICD), Medical Subject Headings (MeSH), and Online Mendelian Inheritance in Man (OMIM). The ontology enables longitudinal comparisons of patient-care data, diagnoses, and treatments, and integration of data across studies. The ontology also helps connect gene and phenotype information related to human disease. DO provides a computable format of inheritable, environmental, and infectious human disease to facilitate the connection of genetic data, clinical data, and symptoms.

The top-level *Disease* class (DOID:4) has 8 direct subclasses, of which *Disease of anatomical entity* (DOID:7) is further divided into 12 direct subclasses (Table 1). A disease class in DO can have more than one parent in the hierarchy; for example *viral hepatitis* (DOID:1884) is a *hepatitis* (DOID:2237) and a *viral infectious disease* (DOID:934)

#### Human Phenotype Ontology

The Human Phenotype Ontology (HPO) provides a controlled terminology to describe the phenotypic features of hereditary, congenital, and acquired diseases; version 2021-04-13 includes 19 618 entities [9; 11]. Initially, the ontology incorporated only simple Mendelian diseases, with about 50 000 annotations connecting the ontology to 4779 diseases in the OMIM database of genetic disorders. HPO now includes features of more than 3400 common disorders, many of which have complex, poly-genetic risk factors.

HPO terms can have more than one parent in the phenotypic hierarchy: for example, *neoplasm of the stomach* (HP:0006753) has parent terms *abnormal stomach morphology* (HP:0002577) and *neoplasm of the gastrointestinal tract* (HP:0007378). Phenotypes can be matched at varying levels of granularity in the

ontology's hierarchy. The term *phenotypic abnormality* (HP:0000118) and its subclasses form the core part of the ontology; the present analysis focused on the 25 direct subclasses of *phenotypic abnormality* (Table 2).

Table 1 – Principal disease categories, identified as subclasses of the top-level DO Disease class, with further breakout by anatomical entity.

| Number | Disease Category                  |
|--------|-----------------------------------|
| D1     | Disease by infectious agent       |
| D2     | Disease of cellular proliferation |
| D3     | Disease of mental health          |
| D4     | Disease of metabolism             |
| D5     | Genetic disease                   |
| D6     | Physical disorder                 |
| D7     | Syndrome                          |
|        | Disease of anatomical entity      |
| D8     | Cardiovascular system disease     |
| D9     | Endocrine system disease          |
| D10    | Gastrointestinal system disease   |
| D11    | Hematopoietic system disease      |
| D12    | Immune system disease             |
| D13    | Integumentary system disease      |
| D14    | Musculoskeletal system disease    |
| D15    | Nervous system disease            |
| D16    | Reproductive system disease       |
| D17    | Respiratory system disease        |
| D18    | Thoracic disease                  |
| D19    | Urinary system disease            |

Table 2 – Phenotypic abnormality categories.

| Number | Phenotypic Abnormality Category                |
|--------|--|
| 1      | Abnormal cellular phenotype                    |
| 2      | Abnormality of blood and blood-forming tissues |
| 3      | Abnormality of connective tissue               |
| 4      | Abnormality of head or neck                    |
| 5      | Abnormality of limbs                           |
| 6      | Abnormality of metabolism/homeostasis          |
| 7      | Abnormality of prenatal development or birth   |
| 8      | Abnormality of the breast                      |
| 9      | Abnormality of the cardiovascular system       |
| 10     | Abnormality of the digestive system            |
| 11     | Abnormality of the ear                         |
| 12     | Abnormality of the endocrine system            |
| 13     | Abnormality of the eye                         |
| 14     | Abnormality of the genitourinary system        |
| 15     | Abnormality of the immune system               |
| 16     | Abnormality of the integument                  |
| 17     | Abnormality of the musculature                 |
| 18     | Abnormality of the nervous system              |
| 19     | Abnormality of the respiratory system          |
| 20     | Abnormality of the skeletal system             |
| 21     | Abnormality of the thoracic cavity             |
| 22     | Abnormality of the voice                       |
| 23     | Constitutional symptom                         |
| 24     | Growth abnormality                             |
| 25     | Neoplasm                                       |

### Radiology Gamuts Ontology

The Radiology Gamuts Ontology (RGO, version 1.0) incorporates knowledge of radiological differential diagnosis: its 16 912 concepts specify diseases (e.g., *cirrhosis*), imaging observations (e.g., *hepatomegaly*), and interventions (e.g., *partial nephrectomy*) [2]. Its subsumption hierarchy—which expresses relationships between more general and more specific concepts—is relatively sparse: RGO has only 1782 subclass-superclass (“is a”) relationships. RGO’s 55 564 causal relationships—expressed as the “may cause” relation and its inverse “may be caused by” relation—encode the relationships between conditions and their imaging manifestations for radiological diagnosis. For example, RGO’s axioms posit that *cirrhosis* may cause 31 conditions including *ascites* and *chylothorax*, and in turn, may be caused by 35 conditions such as *hepatitis*, *cystic fibrosis*, and *Caroli disease* (<https://www.gamuts.net/x/22460>). RGO’s causal relation indicates a tendency, and is less strict than logical implication. The specified causes of a particular finding may not be exhaustive; thus, RGO supports “open-world” inference.

RGO includes common diseases (e.g., *diabetes*) and rare conditions (e.g., *Cruveilhier-Baumgarten syndrome*). The ontology has been mapped to the Orphanet Rare Disease Ontology (ORDO) [6; 10] and to common biomedical vocabularies such as the Radiology Lexicon (RadLex), Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT), and International Classification of Diseases 10th edition (ICD-10) [3]. RGO has been used to aid in radiological differential diagnosis [5], and has been applied to analysis of radiology reports to estimate the frequency of rare diseases [7].

### Integration of ontologies

RGO has been integrated with DO and HPO to create a broader, more general knowledge resource and to enable sharing and reuse of knowledge across domains [4]. The integration involved mapping RGO terms to corresponding terms in DO and HPO using equivalence (*owl:sameAs*) or subclass-superclass (*is\_a*) relations. For the purposes of this analysis, both relations were included. RGO’s mappings to DO and HPO provide the ability to perform hierarchical abstraction over RGO, which can overcome RGO’s relatively sparse subsumption hierarchy. One particular advantage of the integration of these three ontologies is the ability to pose abstract questions that relate diseases and imaging phenotypes, such as, “Which immune system diseases may cause an abnormality of the gastrointestinal system?” (Figure 1).

The integration affords two perspectives on the knowledge of the ontologies. First, by considering high-level classes of DO and HPO—such as *immune system disease* (DOID:2914) and *abnormality of the digestive system* (HP:0025031)—as abstract superclasses of RGO, one can analyze the range of abstract queries over the RGO ontology. Second, one can exploit the integration to identify new causal relationships between DO and HPO classes.

### Analysis

The current study analyzed RGO entities and their causal relationships as categorized by the high-level subclasses of DO and HPO. We defined the principal DO categories as the 7 direct subclasses of the top-level *disease* class (DOID:4), excluding *disease of anatomical entity* (DOID:7), plus the 12 direct subclasses of the latter term. The 19 principal disease categories are shown in Table 1. We defined the principal phenotypic abnormality categories as the 25 direct subclasses of HPO’s *phenotypic abnormality* entity (HP:0000118). Previously conducted analysis had mapped RGO terms to entities in DO and

HPO, and to their “ancestors” (higher-level concepts) within the DO and HPO hierarchies, respectively [4].

We first tallied the number of RGO terms that appeared in both a disease category and a phenotypic abnormality category. Many RGO terms appear in causal axioms as both causes and effects; for example, *cirrhosis* (RGO:22460) may cause *ascites* and itself may be caused by *hepatitis*. Thus, one can consider *cirrhosis* as both a disease and an imaging finding (phenotypic abnormality).

Second, we tallied all mappings between descendants of the DO principal entities, through their causal relationships to RGO terms, to descendants of HPO’s principal entities. Pairs of causally related terms were identified through SQL queries of the relational database that serves as the editorial repository for RGO. RGO entities are mapped to DO and HPO as either an equivalent concept or as a subclass; the current analysis included both relations.

Of RGO’s 55 554 causal links, 14 737 were mapped to pairs of DO and HPO concepts. The number of causally linked pairs was tallied for each disease category and phenotypic abnormality category, as shown in Figure 4. For example, there were 25 distinct pairs of entities where a *gastrointestinal system disease* (category D10) may cause an *abnormality of the genitourinary system*, as itemized in Figure 2. The greatest number of mapped RGO terms related to musculoskeletal conditions, which is reflected in the frequency of pairs involving the disease category *musculoskeletal system disease* (category D14) and the phenotype categories *abnormality of the skeletal system*, *abnormality of limbs*, and (to a lesser extent) *abnormality of the musculature*. There were no causally related pairs involving the *abnormal cellular phenotype* or *abnormality of the voice*; these findings are not surprising in light of the RGO’s focus on radiological diagnosis.

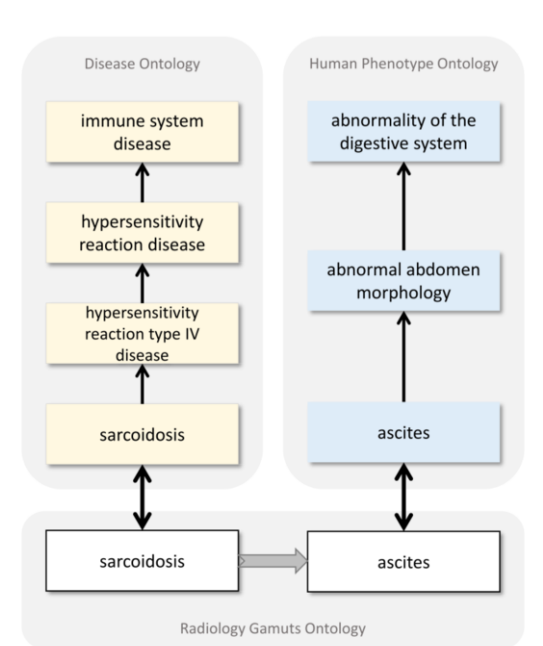


Figure 1 – Relationships across DO, RGO, and HPO show the causal relationship of an immune system disease to a digestive system abnormality. Black single-headed arrows denote subclass (“is\_a”) relations; double-headed arrows denote equivalent terms; thick gray arrows denote causal relations.

## Results

A total of 2275 RGO terms appeared under both a disease category and a phenotypic abnormality category. A detailed analysis is presented in Figure 3. For example, 50 RGO terms were categorized as both a *gastrointestinal system disease* (category D10) and as an *abnormality of the digestive system*.

| Gastrointestinal system disease | Abnormality of the genitourinary system   |
|---------------------------------|---|
| Alagille syndrome               | Nephropathy   |
| Colon cancer                    | Bladder fistula   |
| Colon carcinoma                 | Bladder neoplasm  |
| Crohn's disease                 | Bladder fistula;<br>Perineal fistula;<br>Rectovaginal fistula;<br>Vaginal fistula |
| Diverticulitis                  | Perineal fistula;<br>Rectovaginal fistula;<br>Vaginal fistula                     |
| Diverticulitis of colon         | Bladder fistula   |
| Hepatorenal syndrome            | Nephropathy   |
| Liver cirrhosis                 | Enlarged kidney   |
| Megacolon                       | Vesicoureteral reflux   |
| Peritonitis                     | Renal cortical necrosis   |
| Peutz-Jeghers syndrome          | Ovarian cyst;<br>Ovarian neoplasm   |
| Rectum cancer                   | Rectovaginal fistula;<br>Vaginal fistula  |
| Short bowel syndrome            | Hypercalciuria;<br>Hyperoxaluria  |
| Ulcerative colitis              | Bladder fistula;<br>Rectovaginal fistula  |
| Wilson disease                  | Medullary nephrocalcinosis;<br>Nephropathy  |

Figure 2 – Causally related pairs of gastrointestinal system diseases and phenotypic abnormalities of the genitourinary tract.

Figure 3 – Number of RGO entities that appear within each disease category and phenotypic abnormality category. Disease categories are numbered as in Table 1.

| Phenotypic Abnormality                         | Disease Category |     |    |    |    |    |    |    |    |     |     |     |     |     |     |     |     |     |     |
|--|------------------|-----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|  | D1               | D2  | D3 | D4 | D5 | D6 | D7 | D8 | D9 | D10 | D11 | D12 | D13 | D14 | D15 | D16 | D17 | D18 | D19 |
| Abnormal cellular phenotype                    | 0                | 0   | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| Abnormality of blood and blood-forming tissues | 0                | 13  | 0  | 1  | 1  | 0  | 0  | 3  | 0  | 0   | 25  | 18  | 0   | 1   | 0   | 0   | 0   | 0   | 0   |
| Abnormality of connective tissue               | 0                | 2   | 0  | 0  | 0  | 1  | 0  | 0  | 0  | 7   | 1   | 0   | 4   | 7   | 0   | 0   | 0   | 0   | 0   |
| Abnormality of head or neck                    | 0                | 3   | 0  | 0  | 1  | 4  | 1  | 2  | 0  | 21  | 0   | 1   | 0   | 4   | 11  | 0   | 6   | 0   | 0   |
| Abnormality of limbs                           | 0                | 0   | 0  | 0  | 0  | 5  | 0  | 0  | 0  | 0   | 0   | 0   | 1   | 8   | 0   | 0   | 0   | 0   | 0   |
| Abnormality of metabolism/homeostasis          | 0                | 0   | 0  | 28 | 1  | 0  | 0  | 0  | 1  | 0   | 2   | 4   | 0   | 2   | 1   | 0   | 2   | 0   | 4   |
| Abnormality of prenatal development or birth   | 0                | 1   | 0  | 0  | 0  | 0  | 0  | 2  | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 5   | 0   | 0   | 0   |
| Abnormality of the breast                      | 0                | 2   | 0  | 0  | 0  | 0  | 0  | 0  | 1  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 2   | 0   |
| Abnormality of the cardiovascular system       | 0                | 8   | 0  | 0  | 0  | 1  | 0  | 62 | 0  | 0   | 1   | 1   | 2   | 1   | 2   | 0   | 1   | 0   | 0   |
| Abnormality of the digestive system            | 1                | 14  | 0  | 1  | 1  | 5  | 1  | 2  | 4  | 50  | 0   | 3   | 1   | 0   | 2   | 0   | 0   | 0   | 0   |
| Abnormality of the ear                         | 0                | 1   | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 2   | 0   | 2   | 0   | 0   | 0   | 0   |
| Abnormality of the endocrine system            | 0                | 16  | 0  | 5  | 0  | 0  | 0  | 0  | 36 | 0   | 0   | 3   | 0   | 0   | 0   | 1   | 0   | 0   | 3   |
| Abnormality of the eye                         | 0                | 2   | 0  | 0  | 1  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 1   | 37  | 0   | 0   | 0   | 0   |
| Abnormality of the genitourinary system        | 0                | 17  | 0  | 5  | 2  | 3  | 1  | 2  | 6  | 0   | 3   | 0   | 0   | 1   | 0   | 14  | 0   | 0   | 33  |
| Abnormality of the immune system               | 1                | 8   | 0  | 0  | 0  | 0  | 0  | 4  | 4  | 14  | 3   | 19  | 7   | 6   | 7   | 3   | 8   | 0   | 4   |
| Abnormality of the integument                  | 0                | 8   | 0  | 1  | 4  | 0  | 0  | 1  | 0  | 0   | 0   | 4   | 25  | 3   | 1   | 0   | 0   | 0   | 0   |
| Abnormality of the musculature                 | 0                | 2   | 0  | 1  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 13  | 3   | 0   | 0   | 0   | 0   |
| Abnormality of the nervous system              | 0                | 26  | 10 | 2  | 2  | 10 | 0  | 2  | 12 | 3   | 0   | 0   | 0   | 0   | 50  | 0   | 0   | 0   | 0   |
| Abnormality of the respiratory system          | 0                | 9   | 1  | 0  | 0  | 0  | 0  | 4  | 0  | 1   | 0   | 0   | 0   | 0   | 0   | 0   | 27  | 0   | 0   |
| Abnormality of the skeletal system             | 0                | 8   | 0  | 0  | 3  | 8  | 0  | 0  | 0  | 4   | 0   | 0   | 0   | 59  | 2   | 0   | 1   | 0   | 0   |
| Abnormality of the thoracic cavity             | 0                | 0   | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| Abnormality of the voice                       | 0                | 0   | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 1   | 0   | 0   | 0   | 0   |
| Constitutional symptom                         | 0                | 0   | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| Growth abnormality                             | 0                | 0   | 0  | 1  | 0  | 0  | 0  | 0  | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| Neoplasm                                       | 0                | 130 | 0  | 0  | 1  | 0  | 1  | 3  | 20 | 13  | 5   | 13  | 7   | 17  | 2   | 2   | 4   | 2   | 4   |

Figure 4 – Number of distinct pairs of causally related RGO entities for each disease category and phenotypic abnormality category. Disease categories are numbered as in Table 1.

| Phenotypic Abnormality                         | Disease Category |     |    |     |     |    |     |     |     |     |     |     |     |      |     |     |     |     |     |
|--|------------------|-----|----|-----|-----|----|-----|-----|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|-----|
|  | D1               | D2  | D3 | D4  | D5  | D6 | D7  | D8  | D9  | D10 | D11 | D12 | D13 | D14  | D15 | D16 | D17 | D18 | D19 |
| Abnormal cellular phenotype                    | 0                | 0   | 0  | 0   | 0   | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0    | 0   | 0   | 0   | 0   | 0   |
| Abnormality of blood and blood-forming tissues | 10               | 20  | 0  | 19  | 19  | 0  | 5   | 12  | 7   | 10  | 41  | 33  | 7   | 16   | 3   | 1   | 4   | 0   | 5   |
| Abnormality of connective tissue               | 3                | 3   | 3  | 9   | 39  | 4  | 7   | 2   | 5   | 4   | 0   | 11  | 9   | 26   | 1   | 0   | 0   | 0   | 1   |
| Abnormality of head or neck                    | 11               | 71  | 16 | 86  | 458 | 51 | 173 | 6   | 43  | 18  | 18  | 35  | 29  | 388  | 43  | 0   | 12  | 0   | 3   |
| Abnormality of limbs                           | 12               | 18  | 8  | 87  | 349 | 39 | 88  | 9   | 46  | 10  | 21  | 22  | 25  | 433  | 28  | 0   | 2   | 0   | 8   |
| Abnormality of metabolism/homeostasis          | 14               | 29  | 1  | 31  | 33  | 2  | 12  | 19  | 24  | 17  | 6   | 23  | 3   | 32   | 2   | 1   | 11  | 0   | 32  |
| Abnormality of prenatal development or birth   | 2                | 9   | 0  | 7   | 15  | 6  | 3   | 3   | 4   | 5   | 6   | 3   | 0   | 10   | 1   | 1   | 1   | 0   | 4   |
| Abnormality of the breast                      | 0                | 2   | 2  | 0   | 4   | 0  | 1   | 0   | 8   | 1   | 0   | 1   | 0   | 0    | 0   | 0   | 0   | 1   | 0   |
| Abnormality of the cardiovascular system       | 54               | 48  | 11 | 103 | 134 | 3  | 62  | 248 | 51  | 12  | 37  | 68  | 15  | 106  | 21  | 4   | 14  | 0   | 12  |
| Abnormality of the digestive system            | 79               | 85  | 5  | 113 | 60  | 15 | 36  | 23  | 34  | 128 | 21  | 69  | 12  | 42   | 6   | 16  | 3   | 0   | 13  |
| Abnormality of the ear                         | 1                | 2   | 2  | 14  | 79  | 5  | 24  | 1   | 12  | 1   | 2   | 2   | 3   | 61   | 10  | 0   | 1   | 0   | 3   |
| Abnormality of the endocrine system            | 7                | 14  | 1  | 16  | 48  | 5  | 10  | 1   | 31  | 0   | 2   | 13  | 2   | 9    | 1   | 0   | 0   | 0   | 3   |
| Abnormality of the eye                         | 4                | 22  | 10 | 51  | 197 | 17 | 51  | 8   | 8   | 5   | 9   | 13  | 9   | 102  | 18  | 0   | 4   | 0   | 2   |
| Abnormality of the genitourinary system        | 26               | 68  | 4  | 47  | 127 | 18 | 38  | 16  | 33  | 25  | 14  | 26  | 10  | 27   | 10  | 17  | 1   | 0   | 61  |
| Abnormality of the immune system               | 71               | 54  | 1  | 61  | 22  | 2  | 12  | 12  | 12  | 40  | 22  | 79  | 15  | 27   | 10  | 0   | 24  | 1   | 2   |
| Abnormality of the integument                  | 4                | 7   | 5  | 44  | 172 | 8  | 53  | 10  | 18  | 4   | 4   | 31  | 49  | 43   | 5   | 0   | 0   | 0   | 2   |
| Abnormality of the musculature                 | 4                | 16  | 1  | 18  | 32  | 2  | 10  | 2   | 6   | 4   | 2   | 5   | 2   | 46   | 12  | 0   | 0   | 0   | 0   |
| Abnormality of the nervous system              | 57               | 72  | 14 | 76  | 104 | 18 | 27  | 20  | 24  | 14  | 10  | 27  | 13  | 66   | 82  | 0   | 6   | 1   | 3   |
| Abnormality of the respiratory system          | 64               | 64  | 5  | 10  | 31  | 3  | 14  | 52  | 7   | 23  | 6   | 56  | 4   | 60   | 13  | 3   | 107 | 0   | 14  |
| Abnormality of the skeletal system             | 97               | 176 | 33 | 385 | 790 | 89 | 234 | 44  | 206 | 51  | 68  | 165 | 72  | 1305 | 116 | 0   | 12  | 0   | 32  |
| Abnormality of the thoracic cavity             | 4                | 3   | 1  | 2   | 3   | 0  | 0   | 0   | 0   | 1   | 0   | 3   | 0   | 2    | 0   | 0   | 4   | 0   | 1   |
| Abnormality of the voice                       | 0                | 0   | 0  | 0   | 0   | 0  | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0    | 0   | 0   | 0   | 0   | 0   |
| Constitutional symptom                         | 1                | 2   | 0  | 1   | 0   | 0  | 0   | 4   | 0   | 0   | 2   | 5   | 0   | 7    | 0   | 0   | 0   | 0   | 0   |
| Growth abnormality                             | 0                | 1   | 1  | 2   | 15  | 1  | 6   | 0   | 2   | 0   | 3   | 0   | 1   | 27   | 0   | 0   | 0   | 0   | 1   |
| Neoplasm                                       | 2                | 104 | 1  | 5   | 51  | 6  | 17  | 3   | 10  | 5   | 7   | 15  | 11  | 30   | 1   | 3   | 0   | 0   | 0   |

## Discussion

Of the 16 912 entities in RGO, 12 878 entities (76.1%) act as “causes” (conditions that may cause another entity) and 4662 (27.6%) appear in axioms as “effects”; there are 1206 terms (7.1%) that appear as both. The 2275 RGO entities (13.5%) that had mappings to both DO and HPO concepts indicates the greater number of concepts that are considered as both diseases and phenotypic abnormalities, i.e., imaging findings, even if those entities don’t appear as both causes and effects in RGO axioms. The patterns of overlap are logical: items within a disease category are found primarily in a corresponding phenotypic abnormality category, such as DO’s *disease of cellular proliferation* and HPO’s *neoplasm*.

Analysis of RGO’s identified 14 737 pairs of 55 554 causally related entities (26.5%) that were mapped to pairs of DO and HPO entities. The distribution of pairs provides insights into patterns of disease and their related imaging manifestations.

The current analysis was limited in that it considered only high-level DO and HPO categories. One could perform such an analysis among terms at any levels within the hierarchies of diseases and phenotypic abnormalities. Because both DO and HPO are polyhierarchies—they admit an entity to have more than one parent—pairs of RGO terms may have been counted in more than one category for diseases and phenotypic abnormalities.

The information obtained from this analysis will have value in pedagogical applications and to understand the effects of various diseases across organ systems. The Gamuts Ontology has been used to generate multiple-choice quiz questions, which could be extended using the information obtained here to include questions such as, “Which of the following disease can cause an abnormality of the genitourinary tract?”

## Conclusions

The present analysis provides a valuable perspective on the categorization of Radiology Gamuts Ontology concepts for radiological diagnosis in terms of well-established ontologies for human diseases and phenotypes. The results also provide insights into causal relationships between the Disease Ontology and the Human Phenotype Ontology. Future work will focus on application of this information into pedagogical applications and clinical decision support tools.

## References

- [1] O. Bodenreider, Biomedical ontologies in action: role in knowledge management, data integration and decision support, *Yearb Med Inform* (2008), 67-79.
- [2] J.J. Budovec, C.A. Lam, and C.E. Kahn, Jr., Radiology Gamuts Ontology: differential diagnosis for the Semantic Web, *RadioGraphics* 34 (2014), 254-264.
- [3] R.W. Filice and C.E. Kahn, Jr., Integrating an ontology of radiology differential diagnosis with ICD-10-CM, RadLex, and SNOMED CT, *J Digit Imaging* 32 (2019), 206-210.
- [4] M.T. Finke, R.W. Filice, and C.E. Kahn, Jr., Integrating ontologies of human diseases, phenotypes, and radiological diagnosis, *J Am Med Informatics Assoc* 26 (2019), 149-154.
- [5] C.E. Kahn, Jr., Ontology-based diagnostic decision support in radiology, *Stud Health Technol Inform* 205 (2014), 78-82.
- [6] C.E. Kahn, Jr., Integrating ontologies of rare diseases and radiological diagnosis, *J Am Med Informatics Assoc* 22 (2015), 1164-1168.
- [7] C.E. Kahn, Jr., An ontology-based approach to estimate the frequency of rare diseases in narrative-text radiology reports, *Stud Health Technol Inform* 245 (2017), 896-900.
- [8] W.A. Kibbe, C. Arze, V. Felix, E. Mitraka, E. Bolton, G. Fu, C.J. Mungall, J.X. Binder, J. Malone, D. Vasant, H. Parkinson, and L.M. Schriml, Disease Ontology 2015 update: an expanded and updated database of human diseases for linking biomedical knowledge through disease data, *Nucleic Acids Res* 43 (2015), D1071-1078.
- [9] S. Köhler, S.C. Doelken, C.J. Mungall, S. Bauer, H.V. Firth, I. Bailleul-Forestier, G.C. Black, D.L. Brown, M. Brudno, J. Campbell, D.R. FitzPatrick, J.T. Eppig, A.P. Jackson, K. Freson, M. Girdea, I. Helbig, J.A. Hurst, J. Jahn, L.G. Jackson, A.M. Kelly, D.H. Ledbetter, S. Mansour, C.L. Martin, C. Moss, A. Mumford, W.H. Ouwehand, S.M. Park, E.R. Riggs, R.H. Scott, S. Sisodiya, S. Van Vooren, R.J. Wapner, A.O. Wilkie, C.F. Wright, A.T. Vulto-van Silfhout, N. de Leeuw, B.B. de Vries, N.L. Washington, C.L. Smith, M. Westerfield, P. Schofield, B.J. Ruef, G.V. Gkoutos, M. Haendel, D. Smedley, S.E. Lewis, and P.N. Robinson, The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data, *Nucleic Acids Res* 42 (2014), D966-974.
- [10] A. Rath, A. Olry, F. Dhombres, M.M. Brandt, B. Urbero, and S. Aymé, Representation of rare diseases in health information systems: the Orphanet approach to serve a wide range of end users, *Hum Mutat* 33 (2012), 803-808.
- [11] P.N. Robinson, S. Köhler, S. Bauer, D. Seelow, D. Horn, and S. Mundlos, The Human Phenotype Ontology: a tool for annotating and analyzing human hereditary disease, *Am J Hum Genet* 83 (2008), 610-615.
- [12] L.M. Schriml, C. Arze, S. Nadendla, Y.W. Chang, M. Mazaitis, V. Felix, G. Feng, and W.A. Kibbe, Disease Ontology: a backbone for disease semantic integration, *Nucleic Acids Res* 40 (2012), D940-946.
- [13] S. Schulz and L. Jansen, Formal ontologies in biomedical knowledge representation, *Yearb Med Inform* 8 (2013), 132-146.
- [14] K.C. Wang, Standard lexicons, coding systems and ontologies for interoperability and semantic computation in imaging, *J Digit Imaging* 31 (2018), 353-360.

## Address for correspondence

Charles E. Kahn, Jr., MD, MS, Department of Radiology, University of Pennsylvania, 3400 Spruce Street, Philadelphia, PA 19104 USA.  
email: ckahn@upenn.edu.