Harvard-MIT Division of Health Sciences and Technology
HST.947: Medical Artificial Intelligence
Prof. Peter Szolovits
Prof. Lucila Ohno-Machado

Updating the QMR in 2005: New Approaches

Introduction

Internist-I/Quick Medical Reference (QMR) is a computer-based diagnostic consultant for general internal medicine that was developed by Drs. Miller, Pople, and Myers in the early 1970s.[1] QMR's extensive knowledge base encompasses characteristics of findings and their relationship to diagnoses. Its performance appeared qualitatively similar to that of the hospital clinicians but inferior to that of experts.

We have used the knowledge-acquisition methodology to reconstruct a version of the Internist-I system approach. Our approach requires developing explicit domain ontology, definitions of mappings between the domain ontology and the reusable problem solving method and automated generation of a domain-specific knowledge acquisition tool for entry and editing of content knowledge from various sources including the Research Patients Data Repository (RPDR), the Centers for Disease Control and Prevention (CDC), and Web sites on gene tests.

This project is not an effort to duplicate QMR but to develop an approach to knowledge acquisition and maintenance. We did not test the behavior of resulting system on clinical cases.

Using Ontologies for Internist-I/QMR

Our group used Protégé to reconstruct the well-known QMR/Internist I system to demonstrate the role of a domain ontology - a framework for specification of a model in internal medicine and a reusable problem solving method of updating databases in building a new, workable program.

Protégé is an open platform for ontology modeling and knowledge acquisition. It is a free, open source ontology editor and knowledge-base framework. Protégé is based on Java, is extensible, and provides a foundation for customized knowledge-based applications. The most recent development in standard ontology languages is OWL (Web Ontology Language) from the World Wide Web Consortium. Besides making it possible to describe concepts, OWL has a richer set of operators that make it possible to define concepts as well as to describe them. In addition, the logical model of OWL allows the use of a reasoner that can check if all the statements and definitions in the ontology are mutually consistent and if a concept fits under a given definition. The reasoner is able to check and maintain hierarchy especially when classes have more than one parent. Thus, the OWL Plugin can be used to edit OWL ontologies, to access description logic (DL) reasoners, and to acquire instances for semantic markup. Another plugin is RACER (Renamed ABox and Concept Expression Reasoner). RACER is a Description Logic reasoning system with support for developing ontologies and query answering over RDF documents and with respect to specified RDFS/DAML ontologies.

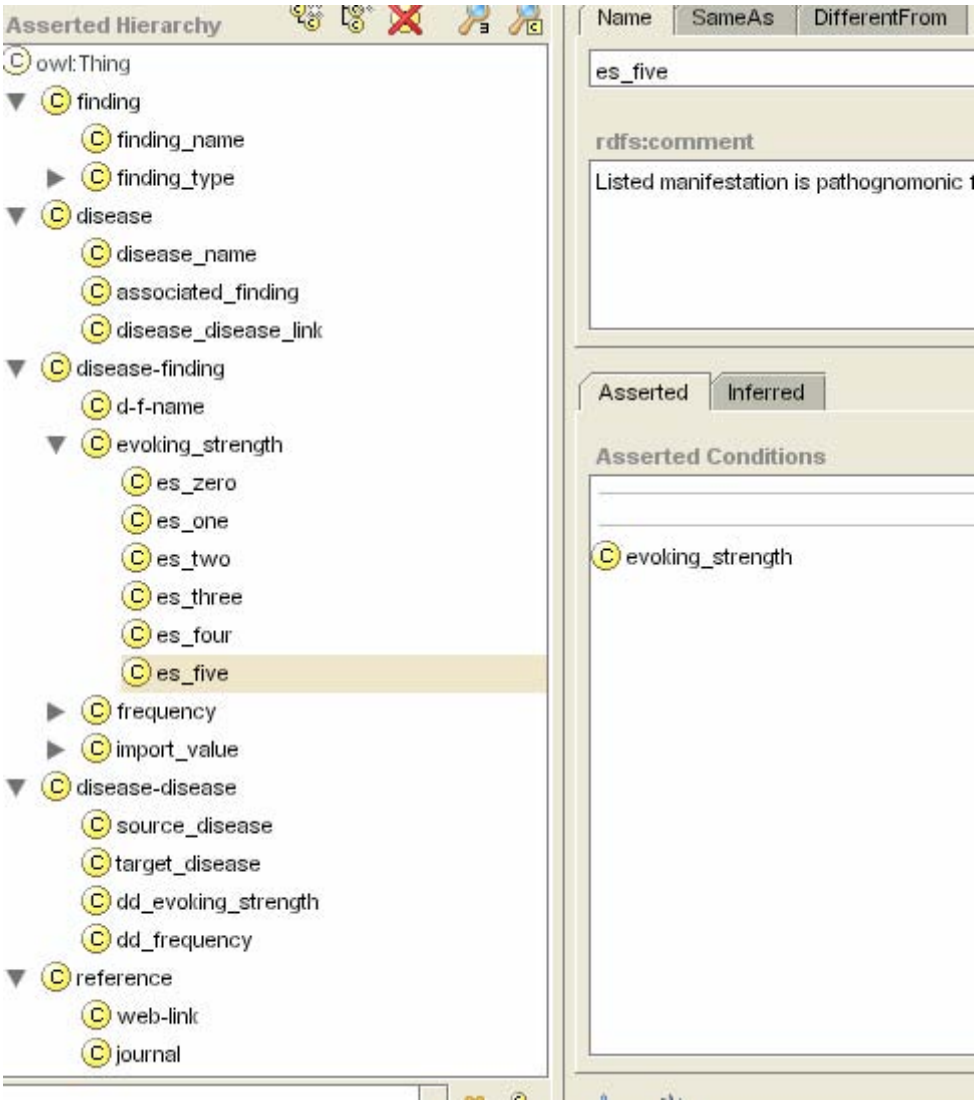
An ontology is a formal, explicit specification of a shared conceptualization. In other words, an ontology describes the concepts in the domain and the relationships that hold between these concepts. It is a shared vocabulary that can be used to model a domain i.e., the objects and /or concepts that exist, their properties and relations. For example, QMR can be represented as a shared vocabulary used to model the domain of internal medicine with the concepts of disease manifestation and findings and the related evoked strengths and frequencies. Ontology can be thought as slightly distinct from knowledge base. Ontology serves a specific purpose of describing the vocabulary and axioms e.g. database schema in QMR is an ontology. Knowledge base includes specific knowledge needed for problem solving. We tried to use additional knowledge bases such as RPDR, CDC, Gene Test and UMLS. The engineering motivation of an ontology is to have a reusable and extendible ontology of the domain. However, to make an explicit ontology is a time consuming developmental process.

There are four important aspects to consider while making an ontology namely: content, form, purpose and development history of the ontology. Content is related to the Object classes and their properties, relationships, and processes while the form of an ontology includes the definition and constraints of the taxonomic relationships, whether the definitional language is as rich as a full logic and whether it is process centric or object-centric. The purpose of an ontology includes knowledge sharing and reuse between people, software systems and agents especially when models or systems change. The development of an ontology is based on factors such as whether it is acquired or

engineered and if acquired what do we know about quality of knowledge, diversity of context, trust in knowledge, and unpredictable use.

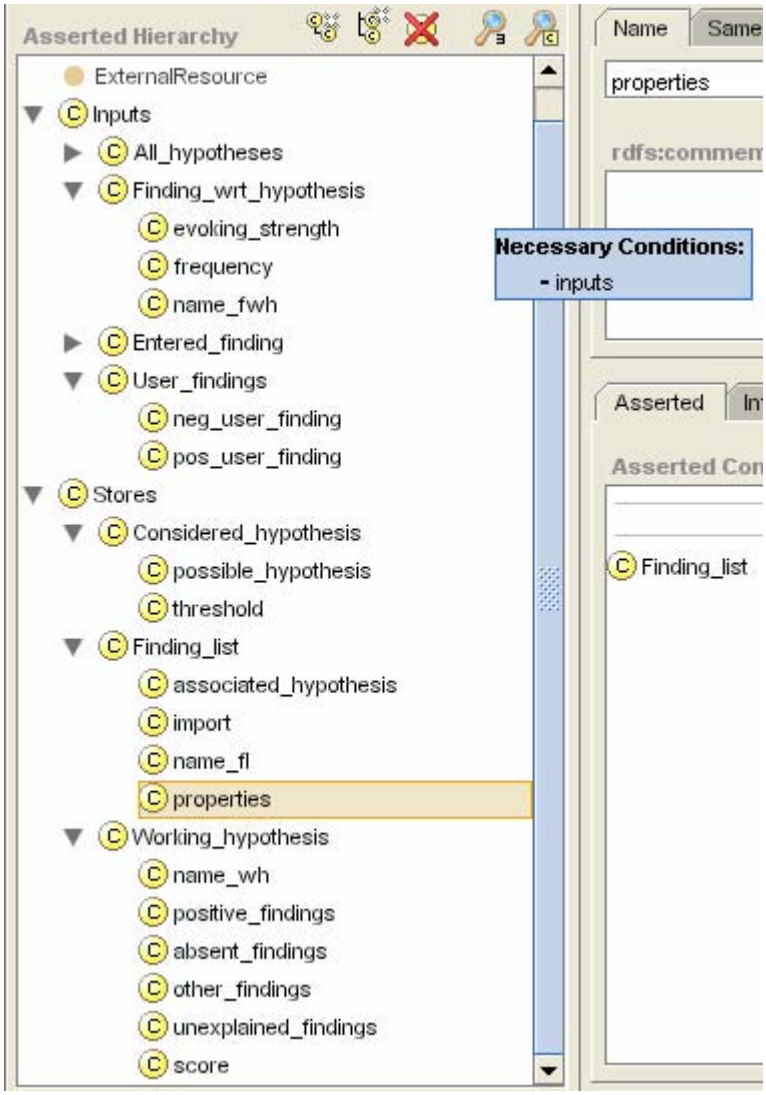
We used Protégé to develop a knowledge-based system to reproduce much of the behavior of INTERNIST-I. We did not have access to a working module of the Internist I. Therefore, our reconstruction is based on the written description of published articles[2] and on a version of QMR made available by Peter Szolovits. Our aim is to demonstrate the use of new approaches to the development of system like Internist that is more explicit and extendible and takes into account the temporal and geographic variance of diseases. Our domain ontology of Internist is relatively simple (Figure1a). There are classes for diseases and findings as well as relationships. One class represents findings whereas another class represents the relations between instances of findings and instances of diseases. The disease-finding class provides a specification of knowledge contained in Internist-I disease profiles. The domain ontology does not contain specific instances of the diseases that can be obtained through automatically updateable knowledge bases.

Figure 1a: Domain Ontology of QMR



The method ontology illustrates the quasi-probabilistic abduction method (Figure 1b) previously illustrated by Mark Musen et al. It defines the inputs to the problem-solving method and the data stores, which are, used internally when the method executes. For example, the working-hypothesis store contains the dynamic list of hypotheses that the problem solving method is considering at any given time. This method is potentially reusable because it is linked to the domain knowledge on which it operates via explicitly mapped relationships. In the case of the Internist-I task, one mapping relation declares that instances of the class "all-hypothesis" in the method ontology are derived from a simple transformation of instances of the "disease" class in the domain ontology; another mapping indicates that instances of "finding-list" in the method ontology simply are "findings" in the domain ontology. The Protégé-II user specifies mappings between the domain ontology and method ontology; a mapping interpreter applies these declarative mappings to the domain knowledge classes and instances so that the problem-solving method accesses the appropriate data elements defined in the method ontology.

Figure 1b: Method Ontology of QMR

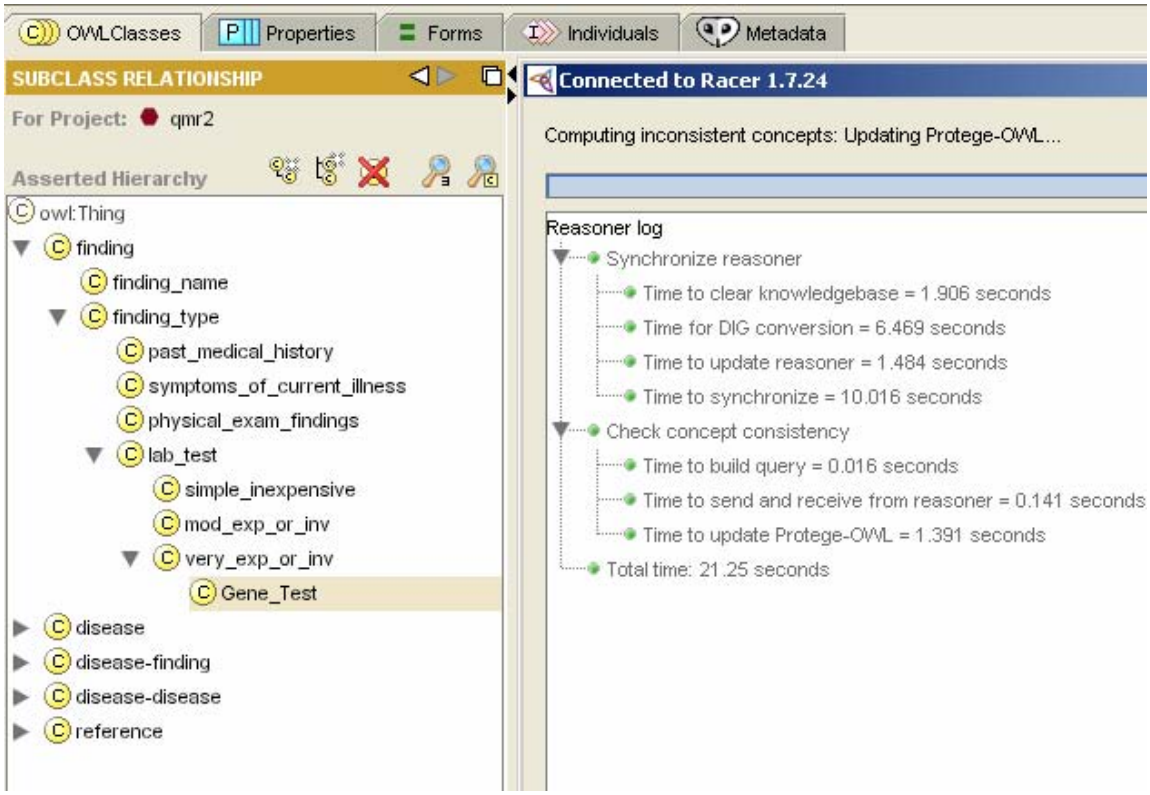


Protégé allows for reusability of problem solving methods and of domain ontologies that traditional knowledge based systems those separate domain knowledge and a reusable inference engine may not. Typical expert-system building shells require the developer to fashion a problem-solving method implicitly from the primitives available in the data elements e.g. production rules on which the inference engine operates. The problem solving method thus becomes inextricably bound up with the same data elements that the developer used to represent domain knowledge.

Using Protégé one can enter description of instances using a domain specific knowledge-acquisition tool that is generated automatically for maintenance of QMR database. RACER provides for a theorem prover within the knowledge acquisition tool to verify that the constraints are not violated by a user's entries. Using RACER one can check for semantic consistency (Figure 2). In addition, additional tabs can now be inserted in Protégé that provide links to UMLS, making sure the ontology has terms consistent and updateable with UMLS.

Therefore, Protégé+OWL+RACER can be used to construct domain ontology for QMR, method ontology and mapping of relationships, tool for knowledge acquisition, updateable tabs for links with other knowledge bases as well consistency checking. The use of explicit domain ontologies, method ontologies and mapping relations in Protégé allows developers to regard domain ontologies and problem-solving methods as well-defined building blocks for the creation of intelligent systems. The construction of explicit mapping relations allows developers to glue-together reusable domain ontologies and problem solving methods when assembling new applications.

Figure 2. Checking Semantic Consistency Using RACER



The Internist-I/Quick Medical Reference (QMR) Inference Engine

Internist-I/QMR knowledge base was first populated in the early 1970s.[1] The current QMR knowledge base includes about 5000 findings, 700 diagnoses, and 53,000 relationships between findings and diagnoses.

Given a set of findings, its inference engine manipulates three basic types of numbers in order to elicit and rank diagnosis hypotheses.[3] The first type of number is the *importance* (IMPORT) of each finding. IMPORTs are a global representation of the clinical importance of findings graded from 1 to 5, with 5 being of highest importance, describing how necessary it is to explain the finding regardless of the final diagnosis. Massive splenomegaly, for instance, has an IMPORT of 5, whereas anorexia has an IMPORT of 1. Mathematical weights are assigned to IMPORT numbers on a non-linear scale.

IMPORT	Description	Examples
1	RARELY require diagnostic consideration	palpitation; dark urine
2	OCCASIONALLY require diagnostic consideration	history of proteinuria
3	USUALLY require diagnostic consideration	oliguria; pica
4	ALMOST ALWAYS require diagnosis explanation	gross hemoptysis
5	MUST ALWAYS be explained diagnostically	Jacksonian seizure; coma

The second type of number is the evoking strength (EVOKS), which describes how strongly one should consider a particular diagnosis versus all other possible diagnoses in the presence of a particular finding. EVOKS is a number that is assigned to a finding/diagnosis pair. A zero indicates that a particular finding is so non-specific that it does not suggest the diagnosis over any other. Again, anorexia is a good example of a non-specific finding. An EVOKS of 5, on the other hand, indicates that the finding is pathognomonic for the diagnosis. Like IMPORT, the EVOKS scale is non-linear.

EVOKS	Description	Finding	Diagnosis
-1	NEVER (TABOS)	white race	sickle cell anemia
0	NONSPECIFIC item(s)	tachycardia	pneumococcal meningitis
1	MINIMALLY SUGGESTS (< 6%) presence of	vertigo	systemic schistosomiasis
2	MILDLY SUGGESTS (6-35%) presence of	hypothermia	acute cardiogenic shock
3	MODERATELY SUGGESTS (36-65%) presence of	dysuria	cystitis
4	STRONGLY SUGGESTS (66-96) presence of	asterixis	hepatic encephalopathy
5	ALWAYS SUGGESTS (> 96%) presence of	hemoglobin SS	sickle cell anemia

(The percentage value in the EVOKS table is the posterior probability of the diagnosis in light of the finding. However, it is unclear how nonspecific and minimally suggests are differentiated. Nonspecific findings should keep posterior probability equal to anterior probability and not necessarily 0.)

The third type of number is frequency (FREQ), which describes the "frequency or incidence of occurrence of a particular clinical finding" in a given disease.[2] Like EVOKS, FREQ is a number assigned to a finding/diagnosis pair. FREQ generally ranges from 1 to 5, with 1 indicating that the finding is rare in the diagnosis and 5 indicating that the finding is present in essentially all instances of the disease. (-1 is used to indicate a

finding that is never found in a diagnosis.) Like IMPORT and EVOKS, the FREQ scale is non-linear.

FREQ	Description	Finding	Diagnosis
-1	NEVER (TABOS)	female sex	pulmonary anthracosis
1	Seen rarely (< 6%) in cases of disease	dyspnea at rest	pyogenic liver abscess
2	Seen in a significant minority (6-35%) of cases	fever	rheumatoid arthritis
3	Seen in about half (36-65%) of cases	history of polydipsia	chronic pyelonephritis
4	Seen in a majority (66-96%) of cases	tachypnea	pneumococcal pneumonia
5	Seen in essentially all (>96%) of cases	myalgia	polymyalgia rheumatica

Each diagnosis is ranked mathematically on the basis of support for it, both positive and negative. The conclusion of a diagnosis is not based on any absolute score, but on how much better is the support for it than for its competitors.

The Limitations of QMR

QMR is remarkable in its breadth of knowledge and capabilities, but it is unable to do the following:

1. Reason anatomically
2. Reason temporally
3. Construct differential diagnoses spanning multiple areas

In order to do the things that it can do, QMR needs the following:

1. A complete and accurate knowledge base.
2. A standardized vocabulary.

QMR can therefore be improved on many fronts. In this design project, we will focus on ways to maintain the knowledge base and keep it up to date. We will also consider the issue of vocabulary.

Updating the Knowledge Base

The QMR inference engine depends on having a complete and accurate knowledge base. To be complete, the knowledge base must contain all possible findings and diagnoses. If a diagnosis is not in the knowledge base, it cannot be concluded. If a finding is not in the knowledge base, it cannot be used as evidence to support or refute diagnoses. To be accurate, IMPORT, EVOKS, and FREQ must have the right values.

Even supposing that the QMR is complete and accurate when it was first created, we still need to maintain it over time to ensure that it remains so. With the passage of time, new findings and diagnoses appear. Conversely, some findings and diagnoses may become obsolete from disuse or replacement. The QMR was developed in the 1970s, and a lot has changed in the practice of medicine since then.

The findings that are used in diagnosis change over time for a number of reasons:

- New imaging technologies create entire classes of new findings. For example, advances in radiology have brought us new sources of findings that are crucial to diagnosis, including computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET). The original QMR database had to be updated to include findings from these technologies.
- Advances in genetics and genomics have also been a source of new types of findings. The results of genetic tests can provide positive or negative support for diagnoses and should be considered.
- New laboratory tests, such as troponin and beta-natriuretic peptide (BNP), supplement or replace older diagnostic tests.

Diagnoses can change over time as well. Some diseases, such as HIV, were first defined after the 1970s. Old diagnoses may be replaced with new ones as medical knowledge advances.

IMPORT, EVOKS, and FREQ values will change over time as well. We need to assign IMPORT values to new findings and EVOKS and FREQ values to new finding/diagnosis pairs. Furthermore, the pattern of old findings in relationship to old diagnoses can change over time as well. It is possible that finding F was commonly associated with diagnosis D in the past, but that is no longer the case now.

To further complicate matters, IMPORT, EVOKS, and FREQ may vary among different populations. Due to differences in genetics and environment, the same diagnosis may manifest differently among different peoples and among different places. Patterns of disease change over time and space. For instance, findings that suggested polio in the past no longer do so now, since polio was eradicated from the United States in 1979 and from the Western Hemisphere in 1991. For an example of geographic variation in disease, consider coccidiomycosis, which is an infection that is endemic in the south-western United States, parts of Mexico and South America.

Therefore, it is not possible to make the QMR knowledge base complete and accurate without specifying when and where the specific instance of QMR will be used. The knowledge base can be made complete and accurate for that time and place, but that same knowledge base may not be accurate for another time and place.

We will describe how we would create as complete and accurate a knowledge base as we can that is applicable in the present and in the Greater Boston area. Other sites that have the same types of data sources that we have can use a similar technique to create their own complete and accurate local version of the knowledge base.

Data Sources

The revised knowledge base will be built on the following:

- The original QMR knowledge base
- Research Patient Data Registry (RPDR)
- Centers for Disease Control and Prevention (CDC)
- Gene Tests (http://www.genetests.org/)

The RPDR is a centralized clinical data registry, or data warehouse, of patient diagnoses, medications, and procedures, used primarily to find research patient cohorts.[4] The RPDR gathers data from various hospital legacy systems and stores it in one place. Information that is available from the RPDR that is of useful for updating the QMR knowledge base include the following:

- Demographics
- Diagnoses
- Laboratory tests
- Medications
- Microbiology
- Procedures
- Transfusion services
- Longitudinal Medical Record data for identified patients
    - Medication list
    - Allergy list
    - Outpatient notes
    - Vital statistics
    - Health maintenance

The data in the RPDR is extensive and comprehensive of all patient encounters from the member hospitals. Its data is therefore representative of the population in the Greater Boston area.

The CDC can provide findings and diagnoses and their relationships for diseases that are reportable. To keep the knowledge base local, we can use the CDC data for locality of interest.

Gene Tests can provide information on the genetic tests available at present.

Using the RPDR

The RPDR is a source for statistics relating findings to diagnoses. For existing QMR finding/diagnosis pairs we can query the QMR to recalculate EVOKS and FREQ and update those values as needed. The RPDR is also a source for new findings and diagnoses. To update the QMR knowledge base, we would look through the findings and diagnoses in RPDR that is not in the QMR and add them to the QMR.

For each new finding, we need an expert or a consensus from experts to assign an IMPORT value. IMPORT is subjective and cannot be determined from the data.

Although we may have the ability to calculate EVOKS and FREQ for every possible finding/diagnosis pair, it would overwhelm the QMR inference engine to have every one of these in the knowledge base. (5000 findings x 700 diagnoses = 3,500,000 possible finding/diagnosis pairs. Adding new findings and diagnoses over time would compound this information explosion.) It makes more sense to restrict finding/diagnosis pairs to those correlations (positive, negative, or neutral) that are known to be common or significant.

For each new finding, determine which clinical presentation would have stimulated the discovery of that finding and what the differential diagnosis of that presentation would be. Then calculate EVOKS and FREQ for each pairing of the finding with a member of the differential diagnosis.

For each new diagnosis, determine what findings are positively or negatively associated with it and calculate EVOKS and FREQ for each pair of finding and the new diagnosis.

An Example of Adding a New Finding to QMR

Beta-natriuretic peptide (BNP) is a blood test that has recently become more popular in the diagnosis of congestive heart failure (CHF). "High BNP" is an example of a new finding that needs to be added to the QMR knowledge base.

First, we ask an expert (or experts) to decide the IMPORT of this finding. Since this is a finding most likely discovered as part of workup in the emergency department when CHF is part of the differential diagnosis, it would make sense to consult an emergency physician and/or a cardiologist.

Second, we determine what diagnoses we should associate the finding with. BNP is ordered when CHF is part of the differential diagnosis, which means that the presentation may include the symptoms of shortness of breath, decreased oxygen saturation, and abnormal breath sounds. The differential diagnosis of this presentation includes pneumonia, pulmonary embolism, acute respiratory distress syndrome, and asthma.

The following tables on the relationship between BNP and members of the differential diagnosis were created using numbers returned by querying the RPDR. The numbers in italics were actual query results. The other numbers were calculated. To keep results up-to-date, the query was restricted to the time period from January 1, 2004, to the present. The population under study is all RPDR patients who had a BNP level measured. In addition, the relationship between BNP and GERD was considered as a control.

	CHF	No CHF	Total
High BNP	*1984*	2009	*3993*
Normal/Low BNP	458	1398	1856
Total	*2442*	3407	*5849*

	PNA	No PNA	Total
High BNP	*718*	3275	*3993*
Normal/Low BNP	234	1622	1856
Total	*952*	4897	*5849*

	PE	No PE	Total
High BNP	*124*	3869	*3993*
Normal/Low BNP	91	1765	1856
Total	*215*	5634	*5849*

	ARDS	No ARDS	Total
High BNP	*104*	3889	*3993*
Normal/Low BNP	23	1833	1856
Total	*127*	5722	*5849*

	Asthma	No Asthma	Total
High BNP	*271*	3722	*3993*
Normal/Low BNP	219	1637	1856
Total	*490*	5359	*5849*

	GERD	No GERD	Total
High BNP	*342*	3651	*3993*
Normal/Low BNP	191	1665	1856
Total	*533*	5316	*5849*

Finding (F)	Diagnosis (D)	Prob(D | F)	EVOKS	Prob(F | D)	FREQ
High BNP	CHF	49.7%	3	81.2%	4
High BNP	PNA	21.9%	2	75.4%	4
High BNP	PE	3.1%	1	57.7%	3
High BNP	ARDS	2.6%	1	81.9%	4
High BNP	asthma	6.8%	2	44.3%	3
High BNP	GERD	8.6%	2	64.2%	3

These results suggest that a high BNP is a finding that suggests CHF, but is not very specific for it.

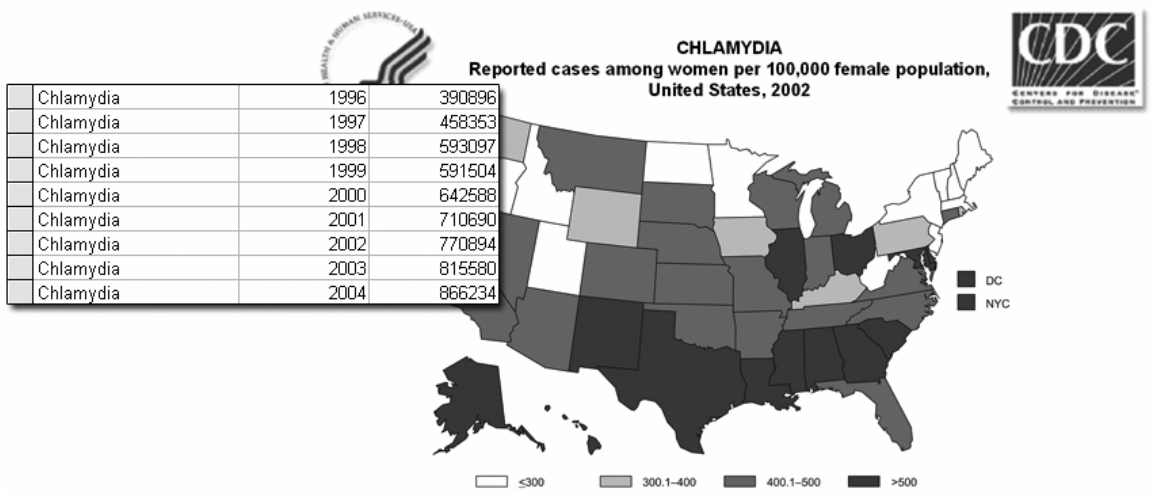
[Note that when we update the QMR, we would not add the finding/diagnosis pair of High BNP/GERD. That pair was just explored as a control.]

Temporal and Regional Trends of Diseases

While QMR utilizes such demographic information as age and gender, it does not address the significant regional differences in disease prevalence nor does it address a mechanism for updating when presented with new temporal disease trend data. An internet-aware and capable means of updating QMR would allow changes to IMPORT and EVOKS values on an arbitrary level of granularity as the disease under investigation may require.

The Centers for Disease Control and Prevention (CDC) distributes a publication entitled the Morbidity and Mortality Weekly Report (MMWR) which provides updates on various diseases and ailments in the population. Included are weekly statistics on a number of reportable diseases. A PERL script was written to extract the information from the MMWR database which was then entered into the QMR Access database.

Information from the MMWR reportable disease statistics includes breakdown of cumulative cases by week along with location of diseases by the entire country, region or state. It is important to note that discussions pertaining to this data are susceptible to reporting practices and current public health programs and efforts. For instance, reported Chlamydia cases have steadily and significantly increased over the past decade (Figure 3). It is presumed that this increase is not due to increased rates of the infection and instead on an expansion in screening activities, improved testing, increased case reporting from providers and improved information systems for reporting. This underscores the importance of utilizing local health information and multiple information sources to help eliminate untoward bias as a result of improved technique which may overestimate disease from an underestimated reference point. For instance, 2000 was the first year in which all 50 states and the District of Columbia instituted regulations requiring Chlamydia reporting thus years prior to 2000 would very significantly underestimate the true burden of disease. Also, overall rates of Chlamydia were highest in the West and Midwest prior to 1996 due to large public resource allocations to screening programs in family clinics and not due to an actual higher rate of Chlamydia infections.

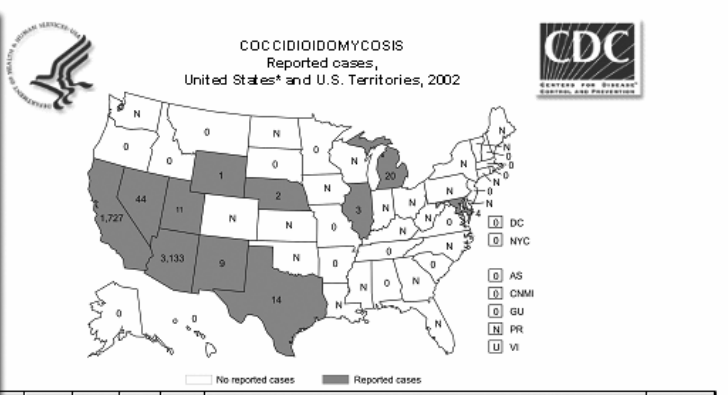


	Chlamydia	1996	390896
	Chlamydia	1997	458353
	Chlamydia	1998	593097
	Chlamydia	1999	591504
	Chlamydia	2000	642588
	Chlamydia	2001	710690
	Chlamydia	2002	770894
	Chlamydia	2003	815580
	Chlamydia	2004	866234

Figure 3 While rates of Chlamydia seem to be increasing as seen from data extracted from the MMWR (left), this is most likely due to improved screening and reporting practices. http://www.cdc.gov/epo/dphsi/annsum/

Despite limitations of reportable diseases due to improved reporting practices and systems over time, such data, especially when used in conjunction with a reliable regional information source, may provide a powerful knowledge base for disease trends. Coccidioidomycosis is a pulmonary disease caused by the inhalation of fungal spores classically described as from the desert regions in the Southwest and characterized by a number of nonspecific findings such as cough, fever, chills, headache, wheezing, loss of appetite and muscle/joint stiffness. This regional criterion is so classic that QMR represents this knowledge in a limited way by a finding of "Residence or Travel Southwestern US Hx" under this disease (Figure 4). However, CDC data more specifically isolates this disease to Arizona and California, which make up over 5,900 of the total 6.056 cases. Thus, an installation of the program that takes this information into account would more correctly assign a frequency of 5 to this finding and more specifically, would account for the current location in that being in Arizona would more highly evoke a diagnosis of this disease than being located in Massachusetts.

Reporting Area	Coccidioidomycosis	
	Cum. 2004	Cum. 2003
UNITED STATES	6,056	4,184
MOUNTAIN	3,932	2,429
Mont.	N	N
Idaho	N	N
Wyo.	2	1
Colo.	N	N
N. Mex.	21	10
Ariz.	3,813	2,374
Utah	37	9
Nev.	59	35
PACIFIC	2,099	1,738
Wash.	N	N
Oreg.	-	-
Calif.	2,099	1,738
Alaska	-	-



			PREV	DX_M	DX_N	MX_IN	FINDING NAME	MX_ID
		Y	2	0	4	1	RACE WHITE <CAUCASIAN>	3128
		Y	2	1	3	3	RALES LOCALIZED	3134
		Y	2	0	4	2	RESIDENCE OR TRAVEL SOUTHWESTERN US HX	3220
		Y	2	1	2	3	RHONCHI LOCALIZED PERSISTENT	3266
		Y	2	1	2	3	RIGOR <S>	3267
		Y	2	0	3	3	SEX FEMALE	3304
		Y	2	0	3	3	SEX MALE	3305
		Y	2	0	2	1	SKIN SWEATING INCREASED GENERALIZED	3493
		Y	2	1	2	3	SPUTUM BLOOD STREAKED	3622
		Y	2	1	1	1	SPUTUM CLEAR MUCOID	3626
		Y	2	2	2	5	SPUTUM CULTURE COCCIDIOIDES	3633
		Y	2	1	2	2	SPUTUM PRODUCTION	3661
		Y	2	1	1	3	SPUTUM PURULENT	3662
		Y	2	2	2	4	SPUTUM SMEAR COCCIDIOIDES	3669
119	COCCIDIOIDOMYCOSIS CHRONIC PULMONARY	Y	2	1	2	5	STOMACH ASPIRATE CULTURE COCCIDIOIDES	3693
119	COCCIDIOIDOMYCOSIS CHRONIC PULMONARY		2	0	2	1	TACHYCARDIA	3762
119	COCCIDIOIDOMYCOSIS CHRONIC PULMONARY		2	0	2	2	TACHYPNEA	3763
119	COCCIDIOIDOMYCOSIS CHRONIC PULMONARY		2	0	1	3	WBC 14000 TO 30000	4025

Figure 4 While QMR (bottom right) somewhat captures localization information with a limited use of findings. However, MMWR data (left) more accurately localizes the disease. http://www.cdc.gov/epo/dphsi/annsum/

Other diseases have less obvious and less recognized yet still apparent geographical trends. The cause of multiple sclerosis, a progressive disease characterized by neurological deficits distributed in time and space, is unknown although it is presumed to have some environmental component due to the predilection of the disease to occur in

northern Europe, northern United States, southern Australia and New Zealand (Figure 5). Using a localized database to modify QMR parameters would help to automatically address this issue of differing geographical trends. In addition, there are numerous diseases and conditions that specifically affect particular populations (sickle cell anemia in the African American population, Tay-Sachs disease in Ashkenazi Jews) which may be addressed not only as specific characteristics of the disease but also region-based information flow.

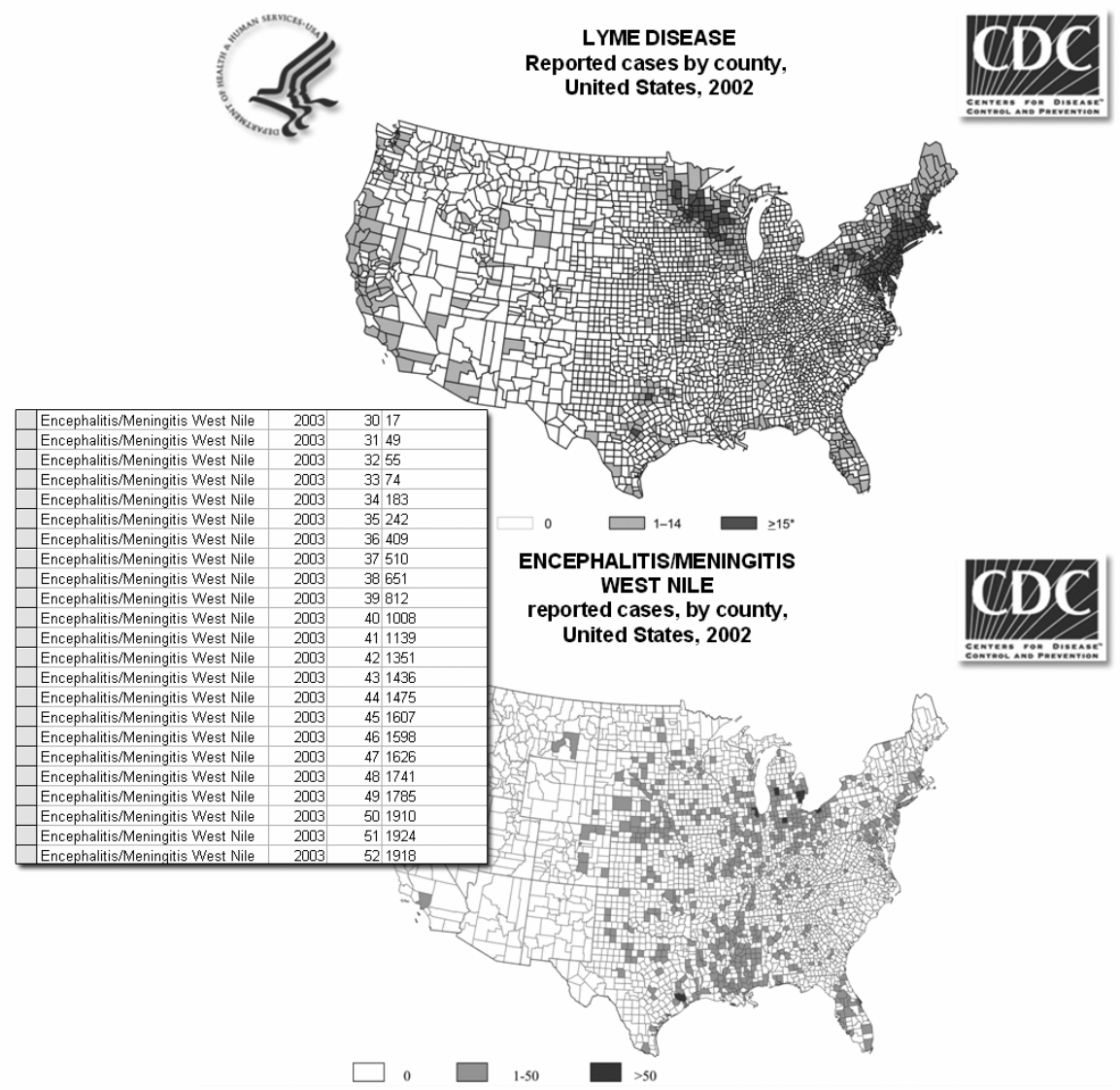
Map removed for copyright reasons. "World Distribution of Multiple Sclerosis." Source: http://medstat.med.utah.edu/kw/ms/mml/ms_worldmap.html

Figure 5 Multiple sclerosis has a definite yet less recognized geographic risk based on latitude that would be automatically captured with a local or regional information source. http://medstat.med.utah.edu/kw/ms/mml/ms_worldmap.html

Temporal trends include gross trends spanning multiple years such as the decline of rubella in the United States and more sporadic trends of rare diseases such as anthrax and plague (Figure 6). Incidents occur that may transiently increase the incidence of a disease and then dissipate soon thereafter. A system should be able to discern such trends as they occur and integrate the occurrence, and possibly an upcoming epidemic, as they happen. For instance, a system should pick up the reporting of anthrax as a result of bioterrorism and respond accordingly to increase the prevalence of that disease in QMR and increase the pertinent parameters as the incidence increases. Likewise, in the 80s, there was an epizootic in prairie dogs and rock squirrels that resulted in increased plague cases in humans for which a system should be able to likewise react in its diagnostic capabilities. A more gradual change is the significant decline of rubella in the United States as a result of vaccination programs. Such programs have a profound effect on disease incidence and will result in a decrease in the prevalence and incidence of disease over time; with access to this data, the prevalence of disease may be updated in QMR along with pertinent findings.

Figure 6 Anthrax (left), plague (center) and rubella (right) are all temporally changing diseases with wildly different rates based on environmental, public health and even political factors. http://www.cdc.gov/epo/dphsi/annsum/

There are also examples of diseases that are distributed both temporally and geographically (Figure 7). For examples, lyme disease, an inflammatory disease spread by the deer tick, may be described as occurring in the Northeast, upper Midwest and along the Pacific coast during the late spring, summer and early fall. Likewise, West Nile virus is transmitted by mosquitoes and may progress to encephalitis or meningitis. Having a vector of the mosquito, the peak occurs in late August and early September when the insects are carrying their highest viral load and then tends to decrease as the weather grows colder and the mosquitoes perish. As may be seen, the number of reported cases of West Nile encephalitis or meningitis starts to increase more sharply at around week 34 and continues into approximately week 42. QMR may be modified to take into account both the geographic variation of diseases such as this, i.e. those carried by vectors that appear seasonally, and the weekly variation of disease as the vector and the vector load wax and wane.



	Encephalitis/Meningitis West Nile	2003	30	17
	Encephalitis/Meningitis West Nile	2003	31	49
	Encephalitis/Meningitis West Nile	2003	32	55
	Encephalitis/Meningitis West Nile	2003	33	74
	Encephalitis/Meningitis West Nile	2003	34	183
	Encephalitis/Meningitis West Nile	2003	35	242
	Encephalitis/Meningitis West Nile	2003	36	409
	Encephalitis/Meningitis West Nile	2003	37	510
	Encephalitis/Meningitis West Nile	2003	38	651
	Encephalitis/Meningitis West Nile	2003	39	812
	Encephalitis/Meningitis West Nile	2003	40	1008
	Encephalitis/Meningitis West Nile	2003	41	1139
	Encephalitis/Meningitis West Nile	2003	42	1351
	Encephalitis/Meningitis West Nile	2003	43	1436
	Encephalitis/Meningitis West Nile	2003	44	1475
	Encephalitis/Meningitis West Nile	2003	45	1607
	Encephalitis/Meningitis West Nile	2003	46	1598
	Encephalitis/Meningitis West Nile	2003	47	1626
	Encephalitis/Meningitis West Nile	2003	48	1741
	Encephalitis/Meningitis West Nile	2003	49	1785
	Encephalitis/Meningitis West Nile	2003	50	1910
	Encephalitis/Meningitis West Nile	2003	51	1924
	Encephalitis/Meningitis West Nile	2003	52	1918

Figure 7 Lyme disease and West Nile virus both vary geographically and temporally based on season due to their vectors. West Nile virus data extracted from MMWR (left) shows its rise in August and September and eventual slowing due to the death of its vector, the mosquito. http://www.cdc.gov/epo/dphsi/annsum/

Genetic Testing

The advent of genetic testing is fostering a new type of medicine; genetests.org currently lists 738 genetic conditions for which tests are available which have been extracted and added to the QMR database. Such knowledge should be integrated into programs such as QMR when applicable. These genetic tests would be similar to other confirmatory tests such as cultures and biopsies and thus would have values similar to these.

Findings/Diagnosis					
DISEASE NAME	PREVAL_ID	DX_MX_EVOKINGSTR	DX_MX_FREQUENCY	MX_IMPORT_ID	FINDING NAME
CRANIAL ARTERITIS	3	4	4	5	ARTERY SUPERFICIAL CRANIAL BIOPSY ARTERITIS
TYPHOID FEVER	1	5	4	5	BONE MARROW BIOPSY CULTURE SALMONELLA TYPHI
TOXOPLASMA MENINGOENCEPHALITIS	2	4	4	5	BRAIN BIOPSY TOXOPLASMA ISOLATION BY ANIMAL INOCCULATION
BRONCHOGENIC CARCINOMA SQUAMOUS CELL TYPE	4	3	4	5	BRONCHOSCOPY ENDOBRONCHIAL BIOPSY MALIGNANT NEOPLASM <NON LYMPHOMA>

While the sensitivity and specificity profiles for the various genetic tests were not available through genetests.org, their values for evoking strength, frequency and importance would be similar including an importance of 5, meaning that the result must be explained by the diagnosis.

Findings/Diagnosis					
DISEASE NAME	PREVAL_ID	DX_MX_EVOKINGSTR	DX_MX_FREQUENCY	MX_IMPORT_ID	FINDING NAME
FAMILIAL MEDITERRANEAN FEVER	1	4	4	5	Familial Mediterranean Fever
HEMOPHILIA A	2	4	4	5	Hemophilia A
HUNTINGTON DISEASE	1	4	4	5	Huntington Disease
POLYCYSTIC LIVER DISEASE	1	4	4	5	Polycystic Liver Disease
PORPHYRIA CUTANEA TARDA	2	4	4	5	Porphyria Cutanea Tarda

UMLS, Standard Vocabularies and QMR

The UMLS provides a powerful substrate on which to expand the vocabulary of QMR. This is important since as medicine progresses, diseases and conditions may have names that become deprecated, replaced by other names that may again later themselves be replaced. A program such as QMR must therefore be manually updated as new terminology arises. Even with such updates, diseases in QMR are still referred to using a single name which itself may be seen as a limitation as many medical conditions may be known and actively called a number of different names. Having links to a mechanism such as the UMLS Metathesaurus CUI, a unique identifier for a particular concept, may provide an expansive vocabulary for such programs that is not only very verbose but also connects to a larger, highly utilized vocabulary source. This metathesaurus, a mechanism to interconnect a variety of medical vocabularies, also allows for standardized means to link both diseases and findings between a variety of systems. QMR already has links recorded between diagnoses and ICD-9 codes, SNOMED-CT and the UMLS CUI; findings are linked only to LOINC and CPT codes.

To demonstrate the richness of vocabulary information in the UMLS Metathesaurus, several examples were selected of obscure or deprecated disease names along with an example of additional possible linkages. The obscure eponym Christmas disease, named after the boy for whom this disease was first described, is actually hemophilia B, the much more common name for the malady. However, QMR has selected the former term to identify the disease instead of the more prominent latter. However, the CUI listed for the disease, C0008533, leads one to discover a rich selection of entries for the concept that not only identifies Christmas disease as hemophilia B, but also recognizes that it is also called Factor IX deficiency, among various other lesser names (Figure 8).

C0008533	L0019070	S0047683	MTHICD9	286.1	Hemophilia B
C0008533	L0019070	S0047683	SNOMEDCT	41788008	Hemophilia B
C0008533	L0019070	S0047683	NCI	C26721	Hemophilia B
C0008533	L0019070	S0047683	NDFRT	C1664	Hemophilia B
C0008533	L0019070	S0047683	MSH	D002836	Hemophilia B
C0008533	L0019070	S0376180	DXP	NOCODE	HEMOPHILIA B
C0008533	L0019070	S1344149	CSP	0438-3499	hemophilia B
C0008533	L0019070	S0899336	SNOMEDCT	41788008	Haemophilia B
C0008533	L0019070	S0899336	MSH	D002836	Haemophilia B
C0008533	L0019070	S2717367	MSH	D002836	Bs, Haemophilia
C0008533	L0019070	S2717368	MSH	D002836	Bs, Hemophilia
C0008533	L0019070	S2720155	MSH	D002836	Haemophilia Bs
C0008533	L0019070	S2720230	MSH	D002836	Hemophilia Bs
C0008533	L0019070	S2729043	MSH	D002836	B, Haemophilia
C0008533	L0019070	S2717045	MSH	D002836	B, Hemophilia
C0008533	L0008533	S0025126	COSTAR	NOCODE	Christmas Disease
C0008533	L0008533	S0025126	MSH	D002836	Christmas Disease
C0008533	L0008533	S0360969	DXP	NOCODE	CHRISTMAS DISEASE
C0008533	L0008533	S0360969	QMR	R0121564	CHRISTMAS DISEASE
C0008533	L0008533	S0363386	CSP	0438-3499	Christmas disease
C0008533	L0008533	S0363386	MTHICD9	286.1	Christmas disease
C0008533	L0008533	S0363386	SNOMEDCT	41788008	Christmas disease
C0008533	L0008533	S0220151	MSH	D002836	Disease, Christmas
C0008533	L0009707	S0217915	ICD9CM	286.1	Congenital factor IX disorder
C0008533	L0015492	S0040220	NCI	C26721	Factor IX Deficiency
C0008533	L0015492	S0040220	MSH	D002836	Factor IX Deficiency
C0008533	L0015492	S1928163	MTHICD9	286.1	Deficiency, factor IX
C0008533	L0015492	S0371316	DXP	U000620	FACTOR IX DEFICIENCY
C0008533	L0015492	S0371316	COSTAR	U000274	FACTOR IX DEFICIENCY
C0008533	L0015492	S0880200	SNOMEDCT	41788008	Factor IX deficiency
C0008533	L0015492	S1345654	CSP	0438-3499	factor IX deficiency
C0008533	L0015492	S0030479	MSH	D002836	Deficiencies, Factor IX
C0008533	L0015492	S0040219	MSH	D002836	Factor IX Deficiencies
C0008533	L0015492	S0030536	MSH	D002836	Deficiency, Factor IX
C0008533	L0019251	S0000535	SNOMEDCT	41788008	Hereditary factor IX deficiency disease
C0008533	L0036897	S0003783	SNOMEDCT	41788008	Sex-linked factor IX deficiency disease
C0008533	L0281495	S0353036	DXP	NOCODE	ANTIHEMOPHILIC FACTOR B DEFICIENCY
C0008533	L0284326	S0353566	DXP	NOCODE	AUTOPROTHROMBIN II DEFICIENCY
C0008533	L0292612	S0393757	DXP	NOCODE	PLASMA THROMBOPLASTIN COMPONENT DEFICIENCY
C0008533	L0292612	S1927605	MTHICD9	286.1	Deficiency, plasma thromboplastin component
C0008533	L0295292	S0394386	DXP	NOCODE	PTC DEFICIENCY
C0008533	L0295292	S0366744	MTHICD9	286.1	Deficiency, PTC
C0008533	L0386625	S0502762	SNOMEDCT	41788008	PTC deficiency disease
C0008533	L0795070	S0845222	SNOMEDCT	41788008	Congenital factor IX deficiency
C0008533	L1708548	S1928159	MTHICD9	286.1	Deficiency, functional factor IX
C0008533	L2871352	S3348592	SNOMEDCT	41788008	Hemophilia B (disorder)
C0008533	L4281405	S4964875	SCTSPA	41788008	hemofilia B
C0008533	L4049786	S4733255	SCTSPA	41788008	deficiencia de factor IX ligada al sexo
C0008533	L4051802	S4735272	SCTSPA	41788008	deficiencia hereditaria de factor IX
C0008533	L4124997	S4808466	SCTSPA	41788008	enfermedad de Christmas
C0008533	L4126601	S4810070	SCTSPA	41788008	enfermedad de deficiencia de PTC
C0008533	L4281406	S4964874	SCTSPA	41788008	hemofilia B (trastorno)

Figure 8 The CUI for Christmas disease leads to the discovery that not only is it hemophilia B, but also Factor IX deficiency.

QMR also refers to the condition known as "Toxemia of Pregnancy," a condition whereby the expecting mother experiences elevated blood pressure, swelling, and protein in her urine, with potential progression to seizures, a condition known as eclampsia. The CUI for this condition, C0032978, is no longer in use in the Metathesaurus since this term is no longer in widespread use in the medical community. Instead, this condition is called pre-eclampsia which has a CUI of C0032914 (Figure 9). While different, the Metathesaurus also has record of the deprecated CUIs with links to the new one. In this case, a separate table in the Metathesaurus correctly linked the two CUIs in a manner

representing that the concept C0032978 now may be found under C0032914. Under this CUI, one may see the older term still present thus older vocabulary is preserved somewhat even after updates. Note that QMR is listed under this CUI as TOXEMIA OF PREGNANCY.

```
C0032914 | L0032914 | S0075818 | MSH          | D011225       | Pre-Eclampsia
C0032914 | L0032914 | S0075818 | NDFRT        | C5048         | Pre-Eclampsia
C0032914 | L0032914 | S0075818 | MTH          | NOCODE        | Pre-Eclampsia
C0032914 | L0032914 | S0397294 | SNOMEDCT     | 398254007     | Pre-eclampsia
C0032914 | L0032914 | S0075809 | MSH          | D011225       | Pre Eclampsia
C0032914 | L0032914 | S1316965 | MTHICD9      | 642.4         | Pre-eclampsia NOS
C0032914 | L0032914 | S1316965 | SNOMEDCT     | 288201007     | Pre-eclampsia NOS
C0032914 | L0013607 | S0077888 | MSH          | D011225       | Proteinuria-Edema-Hypertension Gestosis
C0032914 | L0013607 | S0077886 | MSH          | D011225       | Proteinuria Edema Hypertension Gestosis
C0032914 | L0013607 | S0036401 | MSH          | D011225       | Edema Proteinuria Hypertension Gestosis
C0032914 | L0013607 | S0036409 | MSH          | D011225       | Edema-Proteinuria-Hypertension Gestosis
C0032914 | L0013607 | S0044662 | MSH          | D011225       | Gestosis, Edema-Proteinuria-Hypertension
C0032914 | L0013607 | S0044664 | MSH          | D011225       | Gestosis, Hypertension-Edema-Proteinuria
C0032914 | L0013607 | S0044665 | MSH          | D011225       | Gestosis, Proteinuria-Edema-Hypertension
C0032914 | L0013607 | S0050510 | MSH          | D011225       | Hypertension Edema Proteinuria Gestosis
C0032914 | L0013607 | S0050521 | MSH          | D011225       | Hypertension-Edema-Proteinuria Gestosis
C0032914 | L0014477 | S0038245 | MSH          | D011225       | EPH Complex
C0032914 | L0014478 | S0038249 | MSH          | D011225       | EPH Toxemias
C0032914 | L0014478 | S0038248 | MSH          | D011225       | EPH Toxemia
C0032914 | L0014478 | S0094365 | MSH          | D011225       | Toxemia, EPH
C0032914 | L0014478 | S0094368 | MSH          | D011225       | Toxemias, EPH
C0032914 | L0017506 | S0044663 | NDFRT        | C6500         | Gestosis, EPH
C0032914 | L0017506 | S0044663 | MSH          | D011225       | Gestosis, EPH
C0032914 | L0017506 | S0038247 | MSH          | D011225       | EPH Gestosis
C0032914 | L0040524 | S0747310 | MTHICD9      | 642.4         | Toxemia NOS
C0032914 | L0040524 | S0747310 | SNOMEDCT     | 237280005     | Toxemia NOS
C0032914 | L0040524 | S0747305 | SNOMEDCT     | 237280005     | Toxaemia NOS
C0032914 | L0215744 | S1941811 | CSP          | 2404-7447     | pregnancy toxemia/hypertension
C0032914 | L0315802 | S0075903 | COSTAR       | NOCODE        | Preeclampsia
C0032914 | L0315802 | S0075903 | MEDLINEPLUS  | T1508         | Preeclampsia
C0032914 | L0315802 | S0075903 | MSH          | D011225       | Preeclampsia
C0032914 | L0315802 | S0394119 | DXP          | U001572       | PREECLAMPSIA
C0032914 | L0315802 | S0424243 | CSP          | 4001-0110     | preeclampsia
C0032914 | L0400361 | S0495653 | COSTAR       | U000402       | MATERNAL TOXEMIA
C0032914 | L0566757 | S1316964 | SNOMEDCT     | 398254007     | PE - Pre-eclampsia
C0032914 | L0566773 | S1316971 | SNOMEDCT     | 398254007     | PET - Pre-eclamptic toxemia
C0032914 | L0566773 | S1316970 | SNOMEDCT     | 398254007     | PET - Pre-eclamptic toxaemia
C0032914 | L0566914 | S0636136 | SNOMEDCT     | 398254007     | EPH - Edema, proteinuria and hypertension of pregnancy
C0032914 | L0566914 | S0636138 | SNOMEDCT     | 398254007     | EPH - Oedema, proteinuria and hypertension of pregnancy
C0032914 | L0585173 | S0719533 | SNOMEDCT     | 398254007     | Proteinuric hypertension of pregnancy
C0032914 | L0823479 | S1047408 | ICPC         | W81           | Toxemia (pre)eclampsia
C0032914 | L0823480 | S0995845 | SNOMEDCT     | 288201007     | Pre-eclampsia, unspecified
C0032914 | L0823483 | S1636857 | SNOMEDCT     | 398254007     | Pre-eclamptic toxemia
C0032914 | L0823483 | S1636856 | SNOMEDCT     | 398254007     | Pre-eclamptic toxaemia
C0032914 | L1223231 | S0094369 | MSH          | D011225       | Toxemias, Pregnancy
C0032914 | L1223231 | S0000450 | LCH          | U006029       | Toxemia of pregnancy
C0032914 | L1223231 | S0000450 | SNOMEDCT     | 15394000      | Toxemia of pregnancy
C0032914 | L1223231 | S0075991 | MSH          | D011225       | Pregnancy Toxemia
C0032914 | L1223231 | S0094366 | MSH          | D011225       | Toxemia, Pregnancy
C0032914 | L1223231 | S0408534 | DXP          | NOCODE        | TOXEMIA OF PREGNANCY
C0032914 | L1223231 | S0408534 | QMR          | R0121893      | TOXEMIA OF PREGNANCY
C0032914 | L1223231 | S0408534 | CST          | PREGN TOXEMIA | TOXEMIA OF PREGNANCY
C0032914 | L1223231 | S1646179 | SNOMEDCT     | 15394000      | Toxaemia of pregnancy
C0032914 | L1223231 | S2106737 | CSP          | 2404-7447     | pregnancy toxemia
C0032914 | L1223231 | S4177735 | NCI          | C34943        | Toxemia of Pregnancy
C0032914 | L1223231 | S0075992 | NDFRT        | C5070         | Pregnancy Toxemias
```

Figure 9 Toxemia of Pregnancy is now referred to as Pre-Eclampsia with a new CUI that may be accessed utilizing the deprecated CUI table in the Metathesaurus.

While QMR does not specifically link general findings to the Metathesaurus, this would provide a powerful means to help link QMR to local and regional data sources. The findings that have listings are labs and procedures which use LOINC and CPT, their respective standardized vocabulary used for identification and billing. However, with the addition of SNOMED-CT, there is a wealth of clinical terms that may used for more general findings such as abdominal pain (Figure 10). Abdominal pain has a wealth of variations, ranging from acute to chronic to location, which may be captured using the SNOMED-CT vocabulary and then used as a mechanism through which QMR may be linked to other databases.

```
C0000727 | L0364350 | S3319631 | SNOMEDCT | 9209005   | Acute abdominal pain syndrome
C0232488 | L0276358 | S3228258 | SNOMEDCT | 9991008   | Colicky abdominal pain
C0232488 | L2770029 | S3550595 | SNOMEDCT | 9991008   | Spasmodic abdominal pain
C0232491 | L0276356 | S0363469 | SNOMEDCT | 111985007 | Chronic abdominal pain
C0232491 | L2769601 | S3219262 | SNOMEDCT | 111985007 | Chronic abdominal pain (finding)
C0232492 | L1051481 | S1266048 | SNOMEDCT | 307725002 | [D] Upper abdominal pain
C0232492 | L0276426 | S0411559 | SNOMEDCT | 83132003  | Upper abdominal pain
C0232492 | L2769401 | S3622337 | SNOMEDCT | 307725002 | [D] Upper abdominal pain (context-dependent category)
C0232492 | L2769912 | S3607236 | SNOMEDCT | 83132003  | Upper abdominal pain (finding)
C0232495 | L0276404 | S0384881 | SNOMEDCT | 54586004  | Lower abdominal pain
C0232495 | L2769883 | S3398164 | SNOMEDCT | 54586004  | Lower abdominal pain (finding)
C0344304 | L0486442 | S0892116 | SNOMEDCT | 102614006 | Generalized abdominal pain
C0344304 | L0486442 | S0892064 | SNOMEDCT | 102614006 | Generalised abdominal pain
C0344304 | L0659980 | S0891906 | SNOMEDCT | 102614006 | General abdominal pain-symptom
C0344304 | L2769870 | S3333255 | SNOMEDCT | 102614006 | Generalized abdominal pain (finding)
C0400882 | L1051470 | S1266036 | SNOMEDCT | 235841007 | Chronic nonspecific abdominal pain
C0400882 | L2769600 | S3220281 | SNOMEDCT | 235841007 | Chronic nonspecific abdominal pain (finding)
C0423644 | L0659801 | S0834569 | SNOMEDCT | 162046002 | Central abdominal pain
C0423644 | L2769594 | S3212260 | SNOMEDCT | 162046002 | Central abdominal pain (finding)
C0423646 | L0660108 | S1028741 | SNOMEDCT | 247353003 | Site of abdominal pain
C0423646 | L2769910 | S3544559 | SNOMEDCT | 247353003 | Site of abdominal pain (finding)
C0423650 | L0659811 | S0955602 | SNOMEDCT | 162038003 | Non-colicky abdominal pain
C0423650 | L2769618 | S3436199 | SNOMEDCT | 162038003 | Non-colicky abdominal pain (finding)
C0423651 | L0660067 | S0954989 | SNOMEDCT | 162037008 | No abdominal pain
C0423651 | L2769462 | S3434891 | SNOMEDCT | 162037008 | No abdominal pain (context-dependent category)
C0436922 | L2769447 | S3444690 | SNOMEDCT | 163223001 | On examination - abdominal pain on palpation NOS (context-dependent category)
C0436922 | L2769447 | S3444689 | SNOMEDCT | 268941000 | On examination - abdominal pain on palpation (context-dependent category)
C0436924 | L2769418 | S3444685 | SNOMEDCT | 163214004 | On examination - abdominal pain - right hypochondrium (context-dependent category)
C0436925 | L2769413 | S3444680 | SNOMEDCT | 163215003 | On examination - abdominal pain - epigastrium (context-dependent category)
C0436926 | L2769417 | S3444682 | SNOMEDCT | 163216002 | On examination - abdominal pain - left hypochondrium (context-dependent category)
C0436927 | L2769427 | S3444687 | SNOMEDCT | 163217006 | On examination - abdominal pain - right lumbar (context-dependent category)
C0436929 | L2769426 | S3444684 | SNOMEDCT | 163219009 | On examination - abdominal pain - left lumbar (context-dependent category)
C0436930 | L2769422 | S3444686 | SNOMEDCT | 163220003 | On examination - abdominal pain - right iliac (context-dependent category)
C0436931 | L2769419 | S3444681 | SNOMEDCT | 163221004 | On examination - abdominal pain - hypogastrium (context-dependent category)
C0436932 | L2769420 | S3444683 | SNOMEDCT | 163222006 | On examination - abdominal pain - left iliac (context-dependent category)
C0476308 | L0659499 | S1065899 | SNOMEDCT | 207220009 | [D]Recurrent acute abdominal pain
C0476308 | L0659500 | S1007342 | SNOMEDCT | 271858001 | Recurrent acute abdominal pain
C0476308 | L2768874 | S3623190 | SNOMEDCT | 207220009 | [D]Recurrent acute abdominal pain (context-dependent category)
C0476308 | L2768877 | S3505947 | SNOMEDCT | 271858001 | Recurrent acute abdominal pain (finding)
C0476309 | L0659843 | S1065804 | SNOMEDCT | 207229005 | [D]Other specified abdominal pain
C0476309 | L0660099 | S0924690 | SNOMEDCT | 271859009 | Other specified abdominal pain
C0476309 | L2769398 | S3623068 | SNOMEDCT | 207229005 | [D]Other specified abdominal pain (context-dependent category)
C0476309 | L2769904 | S3458397 | SNOMEDCT | 271859009 | Other specified abdominal pain (finding)
C0478119 | L0659756 | S1079256 | SNOMEDCT | 207589008 | [X]Other and unspecified abdominal pain
C0478119 | L2769463 | S3630944 | SNOMEDCT | 207589008 | [X]Other and unspecified abdominal pain (context-dependent category)
C0522061 | L0660048 | S0932650 | SNOMEDCT | 102613000 | Localized abdominal pain
C0522061 | L0660048 | S3396367 | SNOMEDCT | 102613000 | Localised abdominal pain
C0522061 | L2769881 | S3396480 | SNOMEDCT | 102613000 | Localized abdominal pain (finding)
C0558499 | L1051494 | S1266061 | SNOMEDCT | 274287009 | O/E - abdominal pain
C0558499 | L2769446 | S3444679 | SNOMEDCT | 274287009 | On examination - abdominal pain (context-dependent category)
C0558503 | L2769448 | S3444688 | SNOMEDCT | 163218001 | On examination - abdominal pain - umbilical (context-dependent category)
C0563276 | L1051521 | S1266089 | SNOMEDCT | 285387005 | Left sided abdominal pain
C0563276 | L2769880 | S3390054 | SNOMEDCT | 285387005 | Left sided abdominal pain (finding)
C0563277 | L1051544 | S1266113 | SNOMEDCT | 285388000 | Right sided abdominal pain
C0563277 | L2769909 | S3521668 | SNOMEDCT | 285388000 | Right sided abdominal pain (finding)
C0585107 | L1051543 | S1266112 | SNOMEDCT | 307199009 | Psychosomatic abdominal pain
C0585107 | L2769908 | S3498844 | SNOMEDCT | 307199009 | Psychosomatic abdominal pain (finding)
C0589386 | L1051533 | S1266101 | SNOMEDCT | 304542004 | Nonspecific abdominal pain
C0589386 | L1051480 | S1266047 | SNOMEDCT | 311813008 | [D]Nonspecific abdominal pain
C0589386 | L2769395 | S3622998 | SNOMEDCT | 311813008 | [D]Nonspecific abdominal pain (context-dependent category)
C0589386 | L2769895 | S3437705 | SNOMEDCT | 304542004 | Nonspecific abdominal pain (finding)
C0740577 | L0276335 | S0354326 | SNOMEDCT | 116290004 | Acute abdominal pain
C0740577 | L2768876 | S3319609 | SNOMEDCT | 116290004 | Acute abdominal pain (finding)
C1282002 | L2770030 | S3605843 | SNOMEDCT | 314212008 | Unexplained abdominal pain
C1300119 | L2769619 | S3333276 | SNOMEDCT | 371102005 | Generalized colicky abdominal pain
C1300119 | L2769619 | S3333182 | SNOMEDCT | 371102005 | Generalised colicky abdominal pain
C1300119 | L2769617 | S3333274 | SNOMEDCT | 371102005 | Generalized colicky abdominal pain (finding)
```

Figure 10 Abdominal pain comes in a variety of flavors in SNOMED-CT which provides fertile ground for QMR to link to other databases using the Metathesaurus.

Summary

The QMR can benefit from the application of new methods and new data that have come into being since it was first developed three decades again. We can begin by developing an explicit domain ontology, mapping between the domain ontology and a reusable problem solving method, and creating an automated domain-specific knowledge acquisition tool to gather new knowledge from various sources. Sources of knowledge that we can tap into include the RPDR, the CDC, and genetic testing Web sites.

The knowledge we acquire will update the QMR knowledge base and keep it

complete and accurate by adding findings and diagnoses and updating the EVOKS and FREQ values of finding/diagnosis pairs. Exploring spatiotemporal trends would allow us to customize the QMR to localities and make it more sensitive to epidemics.

Mapping QMR terms to UMLS concepts would standardize the vocabulary in QMR and make it easier for other applications to interact with QMR and to update the knowledge base in a consistent fashion.

---

[1] Miller RA, Pople HE, Myers JD. Internist-1, an experimental computer-based diagnostic consultant for general internal medicine. NEJM 307:468-476, 1982.
[2] Pople HE. Heuristical Methods of Imposing Structure on Ill-Structured Problems: The Structuring of Medical Diagnostics, Artificial Intelligence in Medicine 119-190, 1982.
[3] Myers JD. The Background of INTERNIST-I and QMR. Proc ACM, Bethesda, MD, 195-7, 1987.
[4] Einbinder JS, Murphy SN, Weiner MG. Data warehouses to support clinical research: three approaches. (http://adams.mgh.harvard.edu/PDF_Repository/D010001286.pdf)