



Published in final edited form as:

Stud Health Technol Inform. 2016 ; 225: 432–436.

Towards a HPV Vaccine Knowledgebase for Patient Education Content

Dennis WANG^a, Rachel CUNNINGHAM^b, Julie BOOM^b, Muhammad AMITH^c, Cui TAO^{c,1}

^aUniversity of Texas at Austin

^bImmunization Project, Texas Children's Hospital

^cUniversity of Texas Health Science Center

Abstract

Human papillomavirus is a widespread sexually transmitted infection that can be prevented with vaccination. However, HPV vaccination rates in the United States are disappointingly low. This paper will introduce a patient oriented web ontology intended to provide an interactive way to educate patients about HPV and the HPV vaccine that will empower patients to make the right vaccination decision. The information gathered for this initial draft of the ontology was primarily taken from the Centers for Disease Control and Prevention's Vaccine Information Statements. The ontology currently consists of 160 triples, 141 classes, 52 properties and 55 individuals. For future iterations, we aim to incorporate more information as well as obtain subject matter expert feedback to improve the overall quality of the ontology.

Keywords

Ontology; Vaccine; Human Papillomavirus; Knowledgebase; Patient education

1. Introduction

Human papillomavirus (HPV) is the most common sexually transmitted infection in the United States and currently affects 79 million individuals [3]. There are more than 120 types of HPV [11], some of which are considered high risk and can cause many types of cancers. HPV is spread through the skin and mucous membranes during direct skin-to-skin contact.

Fortunately, the HPV vaccine is an effective tool that can prevent the most common types of HPV infection, which can lead to several types of cancer. Studies have demonstrated that vaccinating young children before they are sexually active produces the strongest immune response and offers the best protection [1]. According to the Centers for Disease Control and Prevention (CDC), HPV vaccine coverage throughout the United States has been dismal

This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License.

¹Cui Tao, PhD, School of Biomedical Informatics, University of Texas Health Science Center, Houston, TX, USA; cui.tao@uth.tmc.edu.

compared to other vaccines recommended at the same age, in large part due to a lack of parental understanding about the disease and vaccine [13].

In today's age of information and technology, patients are offered many options to educate themselves about vaccines. One study indicated that regardless of the accuracy of information, 70% of people turn towards Internet-based sources to retrieve knowledge about vaccines [19]. Other studies have pointed at the prevalence of anti-vaccination content on the web [15; 22]. While it is beneficial for health care providers to communicate to patients directly [20; 27], time constraints can limit and sometimes prohibit these interactions [17; 23]. Most clinics provide printed vaccine educational materials such as brochures, handouts, and booklets for patients to educate themselves [4]. Moreover, federal law requires that healthcare providers give a copy of the current Vaccine Information Statement (VIS) to all patients prior to vaccination. However, the reading levels of parents many times are often below that of the VIS and other patient materials[16; 18; 25], and patients sometimes ignore or do not receive adequate understanding from this kind of inactive learning approach to make critical vaccination decisions [28].

For this project, we propose the use of an ontological knowledgebase that stores patient-focused vaccine information. An ontology can semantically represent domain knowledge in a machine-understandable format, which enables automatic and intelligent queries. Our knowledgebase can represent complex vaccination information and make it scalable and query-able to deliver rich vaccine information as needed. In addition, the information is represented in a patient-friendly format with limited medical jargon so that patients can better understand the information [21]. Moreover, an ontology for patients creates opportunities for more interactive and precise delivery of vaccine information using patient-friendly natural language queries, where patients and parents receive direct and personalized answers, like VAMATA, a "Siri-like" application for online medical assistance in the military setting [26] or for a prototype interactive mobile application for vaccine education [9]. Therefore, we believe that a HPV ontology can serve as a foundation for an application that will empower patients to make vaccination decisions.

The Vaccine Information Statement Ontology (VISO) is an already existing ontology knowledgebase that contains defined classes and relationships pertaining to vaccines and their pathogens [8]. Using the developed VISO framework, our objective is to define HPV knowledge that is machine-readable and semantically-rich so that it can be incorporated into an interactive learning application for patients.

2. Methods

HPV vaccine information was collected from the VISs from the CDC website [1; 2; 6]. Supplemental information was also gathered from Offit and Moser's *Vaccines and Your Child* handbook [24]. Sentences relevant to HPV, and its vaccine were parsed into simple statements called triples, consisting of a subject, predicate, and object. Using the VISO ontology specification, we incorporated the knowledge triples gathered from the sources into the ontology by either using an existing relationship within VISO or creating new ones to better include and represent information. The program we used to construct the ontology is

Protégé 5.0, a commonly acknowledged, open-source software for ontology authoring. The ontology is represented using the Web Ontology language (OWL), which is a standard ontology language. The ontology was evaluated using the Semiotic Evaluation Management System (SEMS). SEMS is a prototype web-based tool that rates an ontology based on different aspects that factor into the overall score [12]. The main qualities that were assessed are *syntactic*, *pragmatic*, and *semantic*. The *syntactic* quality rates the ontology's machine-readability, the *semantic* quality measures the usefulness of the ontology based on the terms' *ambiguity* and *consistency*, and *pragmatic* measures the extensiveness of the ontology. Based on the evaluation result, the ontology can be further improved with several additional iterations.

3. Results

The ontology uses the previously developed VISO representation to define the conceptual class level, but also includes some refinement to accommodate complex HPV vaccine knowledge. The current version of the Vaccine Information Statement Ontology For Human Papillomavirus (VISO for HPV [7], and figure [5]) contains 160 triples, 141 classes (125 subclasses), 52 properties (36 object properties and 16 data properties), and 55 individuals.

We also obtained quality metrics using the semiotic metric evaluation from [10]. Using the SEMS web-based tool, we automatically generated the *syntactic*, *pragmatic*, and *semantic* scores. Table 1 outlines the quality score breakdown.

The overall quality score based on the three quality aspects amounted to 0.73. The *syntactic* score is 0.76 and is comprised of *lawfulness* and *richness* qualities, which are 1.00 and 0.51, respectively. *Semantic* aspect of the ontology scored 0.93. The *interpretability*, *consistency*, and *clarity* qualities make up the *semantic* aspect, which ranks, 0.88, 0.97, and 0.96, respectively. The *pragmatic* aspect score only included one factor, *comprehensiveness*, so its score is 0.50. Currently, the SEMS tool is still under development, and the *accuracy* quality (an aspect of the *pragmatic* quality) could not be calculated.

4. Discussion

Based on the analysis of the scores, our initial HPV ontology proves promising. The *syntactic* score (0.76) demonstrates that the high machine readability of the ontology due to correct use of syntax (*lawfulness*), and lack of total utilization of OWL features (*richness*). However, for this specific ontology and its use-case of only representing vaccine patient knowledge, some of the features may not be necessary, but as we continually develop VISO for HPV, there may be knowledge that may require other OWL features to better model the vaccine information. The *semantic* quality is also relatively high (0.93) resulting from low use of repetitive terms (*consistency*) and low use of ambiguous terms (*clarity*). Overall, the scores are quite promising for initial work, but evaluating the *accuracy* and expanding the use of OWL-based features are needed to fully evaluate the quality.

Creating the ontology posed many challenges. Even though there is already an existing VISO representation from which the HPV ontology is based, a few concept classes and subclasses had to be created or modified in order to better represent some knowledge triples.

For example, a subclass called “*Adjuvant*” was conceptualized to describe specific ingredients added to vaccines, like monophosphoryl lipid A which is added to Cervarix, one of the licensed HPV vaccines. Another example, the subclass “*DNA Virus*” to describe Human Papillomavirus (HPV) was also created to elaborate on the type of virus that the HPV vaccines target. Another difficulty that we experienced was determining whether to include some of the information found pertaining to HPV in the ontology such as side effects, contraindications, and mechanism of action. Ultimately, any information determined to be beneficial and helpful to the patient was included. Additionally, some of the medical terms in the VIS documents are subject to interpretation and could be ambiguous to patients, such as the terms “*mildly ill*” versus “*moderately ill*”. Most patients will be unable to discern the difference between the degrees of illness; therefore, it can become misleading to patients.

By attaining a functional knowledgebase, we can potentially power the intelligence behind patient-centered interactive agents – mobile devices, kiosks, etc. – to improve patient vaccine literacy and address patient questions in which patient-provider interaction is lacking or could be improved. As for the *accuracy* quality, we will need to obtain feedback from subject matter experts to improve the overall quality once the prototype SEMS tool can facilitate *accuracy* evaluation for subject matter experts to rate. We are investigating the possibility of growing the knowledgebase with any information that is lacking; for example, inclusion of additional patient-level education about the cancers caused by HPV, such as cervical and oropharyngeal cancer. While the initial knowledgebase included basic information regarding HPV-related cancers, additional information is needed to improve the robustness of the knowledgebase. Encoding extended information like these into the knowledgebase can help raise awareness for the need of HPV vaccination from a cancer-awareness perspective. A neglected aspect, yet of great interest to us, is to go beyond encoding facts and information about HPV vaccine and consider incorporating multimedia content to enrich the knowledgebase. Ontological knowledgebases could indeed link the various concepts with complementary multimedia content, like video and images, to deliver dynamic information through the patient-centered interactive agents. In addition, there is also evidence that storytelling may have an impact to educate and raise awareness of vaccination [14]. Whether it would be specific stories linked to certain vaccine concepts or a method to semi-automatically generate a story, the implementation of storytelling components in a machine-readable knowledgebase is of interest and will be further explored.

Acknowledgements.

This research is partially supported by the National Library Of Medicine of the National Institutes of Health under Award Number R01LM011829. The authors also gratefully acknowledge the support from the UTHHealth Innovation for Cancer Prevention Research Training Program (Cancer Prevention and Research Institute of Texas grant # RP140103).

References

- [1]. Vaccine Information Statements (VIS): HPV (Human Papillomavirus) Cervarix® VIS, in, Centers for Disease Control and Prevention, 2011.
- [2]. Vaccine Information Statements (VIS): HPV (Human Papillomavirus) Gardasil® VIS, in: Centers for Disease Control and Prevention, ed., Centers for Disease and Control and Prevention, 2013.

- [3]. Genital HPV Infection - CDC Fact Sheet, in: Centers for Disease Control and Prevention, ed., Centers for Disease Control and Prevention, 2014.
- [4]. Vaccine Information Statement: Facts About VISs - Vaccines - CDC, in: Centers for Disease Control and Prevention, ed., Centers for Disease Control and Prevention, 2014.
- [5]. Vaccine Information Statement Ontology (VISO), in, Ontology Research Group, School of Biomedical Informatics, University of Texas Health Science center, 2015.
- [6]. Vaccine Information Statements (VIS): HPV (Human Papillomavirus) Gardasil®–9 VIS, in: Centers for Disease Control and Prevention, ed., Centers for Disease Control and Prevention, 2015.
- [7]. VISO for HPV, in, Bitbucket, <https://bitbucket.org/uthealthontology/viso-for-hpv>, 2015.
- [8]. Amith M, Gong Y, Cunningham R, Boom J, and Tao C, Developing VISO: Vaccine Information Statement Ontology for patient education, *Journal of Biomedical Semantics* 6 (2015), 23. [PubMed: 25973167]
- [9]. Amith M and Tao C, A Prototype Mobile Virtual Assistant for Semantic-Based Vaccine Information Retrieval, in: *Smart Health*, Zheng X, Zeng D, Chen H, Zhang Y, Xing C, and Neill D, eds., Springer International Publishing, 2014, pp. 200–205.
- [10]. Amith M and Tao C, A Web Application Towards Semiotic-based Evaluation of Biomedical Ontologies, in: *The 14th International Semantic Web Conference (ISWC 2015)*, Bethlehem, United States, 2015.
- [11]. Atkinson W, Wolfe C, and Hamborsky J, Centers for Disease Control and Prevention Epidemiology and Prevention of Vaccine-Preventable Diseases, Public Health Foundation, Washington DC, 2012.
- [12]. Burton-Jones A, Storey VC, Sugumaran V, and Ahluwalia P, A semiotic metrics suite for assessing the quality of ontologies, *Data & Knowledge Engineering* 55 (2005), 84–102.
- [13]. Cunningham R and Boom J, Vaccine-Preventable Disease Human Papillomavirus (HPV), in: Texas Children's Hospital, ed.
- [14]. Cunningham RM and Boom JA, Telling stories of vaccine-preventable diseases: Why it works, *SD Med* (2013), 21–26.
- [15]. Davies P, Chapman S, and Leask J, Antivaccination activists on the world wide web, *Archives of Disease in Childhood* 87 (2002), 22–25. [PubMed: 12089115]
- [16]. Davis TC, Bocchini JA, Fredrickson D, Arnold C, Mayeaux EJ, Murphy PW, Jackson RH, Hanna N, and Paterson M, Parent Comprehension of Polio Vaccine Information Pamphlets, *Pediatrics* 97 (1996), 804–810. [PubMed: 8657518]
- [17]. Davis TC, Fredrickson DD, Arnold CL, Cross JT, Humiston SG, Green KW, and Bocchini JA, Childhood Vaccine Risk/Benefit Communication in Private Practice Office Settings: A National Survey, *Pediatrics* 107 (2001), e17–e17. [PubMed: 11158491]
- [18]. Davis TC, Mayeaux EJ, Fredrickson D, Bocchini JA, Jackson RH, and Murphy PW, Reading ability of parents compared with reading level of pediatric patient education materials, *Pediatrics* 93 (1994), 460–468. [PubMed: 8115206]
- [19]. Fox S and Rainie L, The online health care revolution: How the Web helps Americans take better care of themselves, in: *Pew Internet & American Life Project: Online life report*, Pew Research Center, 2000.
- [20]. Gust DA, Darling N, Kennedy A, and Schwartz B, Parents With Doubts About Vaccines: Which Vaccines and Reasons Why, *Pediatrics* 122 (2008), 718–725. [PubMed: 18829793]
- [21]. Jacobson TA, Thomas DM, Morton F, Offutt G, Shevlin J, and Ray S, Use of a low-literacy patient education tool to enhance pneumococcal vaccination rates: A randomized controlled trial, *JAMA* 282 (1999), 646–650. [PubMed: 10517717]
- [22]. Kata A, Anti-vaccine activists, Web 2.0, and the postmodern paradigm – An overview of tactics and tropes used online by the anti-vaccination movement, *Vaccine* 30 (2012), 3778–3789. [PubMed: 22172504]
- [23]. LeBaron C, Rodewald L, and Humiston S, How much time is spent on well-child care and vaccinations?, *Archives of Pediatrics & Adolescent Medicine* 153 (1999), 1154–1159.
- [24]. Offit PA and Moser CA, *Vaccines & your child: separating fact from fiction*, Columbia University Press, 2011.

- [25]. Powers RD, Emergency department patient literacy and the readability of patient-directed materials, *Annals of Emergency Medicine* 17 (1988), 124–126. [PubMed: 3337429]
- [26]. Rodger J and George J, Mobile Speech and the Armed Services: Making a Case for Adding Siri-like Features to VAMTA (Voice-Activated Medical Tracking Application), in: *Mobile Speech and Advanced Natural Language Solutions*, Neustein A and Markowitz JA, eds., Springer New York, 2013, pp. 319–332.
- [27]. Smith PJ, Kennedy AM, Wooten K, Gust DA, and Pickering LK, Association Between Health Care Providers' Influence on Parents Who Have Concerns About Vaccine Safety and Vaccination Coverage, *Pediatrics* 118 (2006), e1287–e1292. [PubMed: 17079529]
- [28]. Wilson FL, Brown DL, and Stephens-Ferris M, Can Easy-to-Read Immunization Information Increase Knowledge in Urban Low-Income Mothers?, *Journal of Pediatric Nursing* 21 (2006), 4–12. [PubMed: 16428009]

Table 1.

Ontology evaluation scores.

	Quality	Quality Score
Syntactic		0.76
	Lawfulness	1.00
	Richness	0.51
Semantic		0.93
	Interpretability	0.88
	Consistency	0.97
	Clarity	0.98
Pragmatic		0.50
	Comprehensiveness	0.50
Overall Score		0.73