A Student Authored Online Medical Education Textbook: Editing Patterns and Content Evaluation of a Medical Student Wiki

Thompson, CL, BA[1], Wade L. Schulz1, BS[1], Adam, Terrence, MD, PhD[2,3].
1.College of Medicine, 2.College of Pharmacy, 3. Institute for Health Informatics, University of Minnesota, Minneapolis, MN, USA

Abstract: The University of Minnesota medical student wiki (UMMedWiki) allows students to collaboratively edit classroom notes to support medical education. Since 2007, UMMedWiki has grown to include 1,591 articles that have collectively received 1.2 million pageviews. Although small-scale wikis have become increasingly important, little is known about their dynamics compared to large wikis, such as Wikipedia. To better understand UMMedWiki's management and its potential reproducibility at other medical schools, we used an edit log with 28,000 entries to evaluate the behavior of its student editors. The development of tools to survey UMMedwiki allows for quality comparisons that improve both the wiki and the curriculum itself. We completed a content survey by comparing the UMMedWiki with two types of rubric data: TIME, a medical education taxonomy consisting of 1500 terms and national epidemiological data on 2,100 diseases.

1) Introduction

Since its inception four years ago, we show that the student-initiated wiki at the University of Minnesota (UMMedWiki) has grown to include almost 1600 articles, which collectively have received nearly 1.2 million pageviews. Although these numbers suggest that the UMMedWiki has become an important asset, little is known about the development of the wiki, its usage, educational content, and its reproducibility at other schools. Currently, there are no techniques available to survey the didactic content of UMMedWiki.

Researchers have investigated the distribution of content creation in Wikipedia and found that 1% of editors contribute 50% of the total edits[1]. This implies that small schools may not have the prerequisite critical mass of student editors that would enable a self-sustaining educational wiki. However, less is known about the dynamics of smaller wikis with specific applications[2,3,4] and what editor base would be necessary for support.

Medical education is complex, requiring the management of scores of expert lecturers. UMMedWiki represents a novel dataset of didactic content for the University of Minnesota that may further elucidate its curricula. When medical schools analyze their own curricula, they rely on evaluation systems such as CurrMitt[8], Knowledge Map[9], or management systems like those at Tufts[10] or the University of Hawaii.[11] Many of these systems are labor-intensive, and all of them organize the curricula by the attribution of metadata to heterogeneous curricular documents such as course catalogues, tests, or PowerPoint slides. Other methods have included the polling of academic deans[12] and manually tabulating textbook entries.[13] Since the 1600 pages of lecture notes in UMMedWiki provide a homogeneous, comprehensive and up-to-date set of curricular documents, the wiki has the potential to more accurately map the curricula than any existing system.

Researchers have analyzed Wikipedia for its content coverage of topics relevant to medical education, specifically pharmacology,[5] pathology informatics[6], and medical informatics.[7] Although several medical schools utilize wikis to support education, very little information is available characterizing the content of these systems.

Given the need to assess the medical student wiki's value, its reproducibility at other medical schools, and its potential role as a curricular mapping system, we conducted research in three broad areas: The first included an exploration of the wiki's development. Specifically we collected an account of the founding of the wiki, the rationale for its creation, and the policies that support its use within the medical school. Second we outlined the usage of UMMedWiki by quantifying the wiki's viewership, its editors, and their relative contributions. Considering UMMedWiki's distinct origin and small scale, we posited that its editing dynamics would markedly diverge from Wikipedia's.

The third area of research became the development of methods to survey contents of UMMedWiki's so that we could better characterize the wiki's organization, didactic coverage, and the relative emphasis given to various concepts. We assumed that rare diseases would be less frequently cited compared to common diseases. Implicitly, UMMedWiki is a reflection of the actual curriculum since students depend on the lecture notes contained within as they prepare for tests, and otherwise conduct studies. With future research we hope to explicitly characterize the relationship between the actual curriculum and the wiki lecture notes.

Research Approach: In summary, this project addresses three main areas concerning UMMedWiki: development, usage, and didactic content. The development of UMMedWiki occurred via evolution from a paper-based note-taking collective to a popular online collaborative system. This development has important implications for wiki implementation: the original paper-based collective insisted on accountability and equitable distribution of work. This ethic may have continued as students transitioned to the wiki-system, which theoretically positively affects the quality of the wiki.

The assessment of UMMedWiki usage is comprised of of two parts: viewership and editing behavior. We focus on editing due to its important implications for the accuracy and thoroughness of content. We examined the metadata recorded by MediaWiki software as editing occurs since this information directly affects metrics used for quality analysis. Editor diversity—defined as the absolute number of unique editors-- positively influences the value of an article.[14]

Finally, to map the didactic content of UMMedWiki, we utilized two rubrics: an epidemiological dataset containing 2100 diseases and an open-source medical education terminology. We visualized the results of this survey to clearly identify areas of disparate content representation.

2) Background

2.1) Development: the Rise of Wikis

In general, the term "wiki" refers to a website that allows for the creating and editing of any number of interlinked web pages via a web browser with a user-friendly interface, accessible to any number of collaborators. Given the flexibility of wikis, their applications and development have been diverse: Wikipedia began as process for soliciting openly edited articles that were then selected for a formal review process akin to traditional encyclopedias. The Medical School at the University of California at San Francisco utilized a privately developed, corporate-style wiki called Confluence. According to staff at UCSF, the administration implemented the wiki at the behest of students, who reference it during student-run tutorial sessions. Stanford Medical School's wiki primarily consists of a "student life" guide. Meanwhile, UMMedWiki runs MediaWiki, the same open-source software that supports Wikipedia. Each type of software has its own cost-benefit profile; we note that for research purposes, it may be helpful to run identical software to Wikipedia, the best-researched wiki.

2.2) Usage: Editing Dynamics in Wikipedia

Intuitively, broad participation in editing a wiki would favorably impact its quality; a larger expert pool would be expected to contribute clarifying content, upload more images, add hyperlinks and generally create more value. In Wikipedia, higher quality *is* correlated to the number of editors that have contributed to an article. The community itself defines "quality" through a nomination process; articles deemed high quality become "Featured Articles" and eligible for showcase on Wikipedia's homepage. Remarkably, an algorithm can distinguish Wikipedia's Featured Articles from an average set of articles without data regarding the nomination process. The assessment is based on meta-data alone, such as the number of participating editors, article length, total number of edits, and number of embedded hyperlinks.[15]

Editors in Wikipedia or UMMedWiki create a valuable resource without the expectation of pay; in economics literature, these editors would be called the "k-group" which designates a group that voluntarily provisions a public good. Data is scarce on the size of Wikipedia's k-group, defined here as *the proportion of editors to viewers*. One survey from 2010, found 14% of readers also made edits. However only 3% reported editing more than once every few weeks. Since editors may derive some satisfaction from wide viewership, when viewership

diminishes, the k-group may shrink or reduce its productivity. Through successive blocks of Chinese Wikipedia, the Chinese government has manipulated the audience available to contributors outside of Mainland China. When their viewership is reduced, contributions by existing editors outside Mainland China decline by 41%.

The k-group itself is heterogeneous; with Wikipedia some editors may contribute thousands of edits, while others only a few hundred. Calculating the exact distribution of work in Wikipedia has been controversial. Jimmy Wales, the founder of Wikipedia, argued that most of the work on Wikipedia is done by a small number of editors, "citing that as of December 2004, 2.5% of the registered [editors] on the site made half of the edits."[16] Chi found that 1% of the editors were contributing about 50% of the edits around 2002, but this fell to only 20% by mid-2006 as the ranks of less prolific editors grew. Others have described the distribution of edits as following a power law (following the form $P(r)= 1/r^k$); where "*k*", averages 1.14.[17] This variable measures the disparity in contributions among editors. Since a *lower* value corresponds to less disparity between editors, this represents a *higher* degree of egalitarianism. Recently, Preidhorsky *et al.* assessed "value" by factoring in the longevity of editor contributions, as well as the pageviews of those contributions. This revealed that within Wikipedia, the top 0.1% of editors contributed 44% of the content that users see when they visit Wikipedia.[18]

We hypothesized the UMMedWiki editors would be more egalitarian—with evenly distributed editing—given the wiki's small size and origins as a note-taking collective. While Priedhorsky's sophisticated methods allow researchers to investigate the contribution of value, we investigated a simpler proxy: the contribution of edits.

Wikipedia researchers have tried to assess the influence of "malevolent editors". Because vandals sometimes destroy or alter content, there is concern about the validity of content in Wikipedia. Despite a famous paper detailing Wikipedia's reliability compared to Encyclopedia Britannica,[28] the concern remains that readers of Wikipedia may view damaged material. The primary method for detecting damage—including erroneous content, or deletions of material—has become identifying "reverted" material. Editors in Wikipedia can always turn back the clock by reverting to an earlier version of an article; this is often easier than rewriting damaged content. Up to 6% of edits in Wikipedia constitute damage repair.[25]

2.4) Content: Making a Map

Researchers have studied medical education's coverage of chronic fatigue,[13] gerontology,[19] or women's health issues.[20] Using computation, Denny *et al.* surveyed seven broad categories of infectious disease, including bacteria, viruses, fungi, and prions. Denny found that "prions" were over-represented in the curricula compared to clinical practice by a factor of 40:1. His method used natural language processing to compare 2,048 curricular documents such as PowerPoint slides to 330,000 student-authored clinical notes.[21]

Computational tools are also used at medical schools evaluate their educational content. One computational tool, CurrMIT, is a relational database, which facilitates tagging of entries with standardized keywords. The medical school at Vanderbilt uses Knowledge Map[9] that employs natural language processing to classify curricular documents. In addition to computational tools, there are medical education taxonomies available to map medical school curricula. In Scotland, Haig *et al.* created the Medical Education Taxonomy Research Organization (METRO) to support education.[22] Willet responded to a perceived lack of "content terms" in METRO by developing an open-source taxonomy called (TIME) that includes 1,500 concepts and can be used as an "index for curriculum mapping, meta-tagging learning objects, or categorizing examination questions."[15] Overall, TIME's open source license made this the most accessible rubric for a survey of UMMedWiki for this study.

3) Methods

3.1) Development:

Much of this section, which reports the origin and supporting policies of UMMedWiki, have become common knowledge at the University of Minnesota Medical School. Here we drew primarily on the experience of this paper's student authors.

3.2) Usage: Quantifying Editors and Viewers

To address the question of viewership of UMMedWiki, data regarding visitor traffic on UMMedWiki were derived from Google Analytics. A Google analytics account was used to monitor usage of UMMedWiki starting in

2007, before its public debut. Using this data, we gathered statistics for the number of pageviews, unique pageviews, and the time users spent on the site.

To characterize the behavior of the editors of UMMedWiki, we used features embedded in the MediaWiki software that supports UMMedWiki. MediaWiki software records changes to the wiki in the "User Edit Log"; this provides information on edit timing, the screen identity of editors, and the types of edits completed including page creation, edits and damage repair (such as "reverted pages"). Since MediaWiki records the identity of editors, we also looked at "editor diversity" over the four years of UMMedWiki operation. This is a measure of participation, because it looks at how many different students edited the wiki at a given point in time.

3.3) Content: Cataloguing the UMMedWiki's Educational Content

In order ascertain how much of UMMedWiki is devoted to topics related to disease concepts versus other concepts; we modified TIME by labeling concepts that matched concepts in SNOMED (January 27, 2011 release) defined as a "disorder". This disorder label belongs to the "concept" attribute in SNOMED and is linked by an "IS-A" relationship to parent concepts containing the term "disorder" within SNOMED. Put simply, it is possible to distinguish disorder concepts (aka diseases) in SNOMED from other types of concepts, such as "procedures". We used TIME to query SNOMED since many identical terms occur in the two terminologies; when a matching concept was discovered that also contained the "disorder" label, this meta-data was transferred to the identical concept in TIME, thus enriching TIME without detriment to its normal operation. Essentially, all concepts in TIME commonly recognized as diseases, were labeled as such by referencing SNOMED. Then, after removing TIME concepts known to generate null results (specifically terms that included "Canada" or " / "), we queried UMMedWiki with the remaining 1,357 terms from the TIME Terminology. Matches included strings with variable word order, intervening words, and characters concatenated to the query string.

Given the utility of UMMedWiki as curricular dataset, we gauged the inclusion and relative emphasis of the disease concepts described within the wiki's lecture notes. Our intent became correlating the relative emphasis of educational content with the burden of disease within the population as measured by an epidemiological dataset containing 2100 diseases, which we assembled to serve as a rubric. The wiki itself is structured much like a textbook with chapters corresponding to separate lectures—it does not contain any overview of the hundreds of (or thousands of) diseases discussed throughout its 1600 pages of notes. While we survey UMMedWiki as a particular curriculum, we also show how curriculum mapping might take place within a wiki supporting education.

Although the wiki contains miscellaneous content in addition to lecture summaries (such as material related to student council proceedings mentioned above) this additional content has no explicit label and cannot be excluded from our analysis. Clearly, the wiki was designed to record lecture material, so this comprises the vast majority of content.

Lecture summaries within UMMedWiki typically focus on specific diseases; this is especially apparent by the second year of medical school, which emphasizes pathology. Assessing the coverage of specific diseases within UMMedWiki began with the creation of an epidemiological dataset that required several stages of refinement. The epidemiological dataset consisted of 2,000 uncommon diseases ranked by *prevalence*. We also included *incidence* data on the most frequent 100 hospital problems, as ranked by Agency for Healthcare Research and Quality (AHRQ) hospitalization data. This epidemiological dataset provides an objective measure of relative clinical content. In order to assess the relative content of the UMMedWiki to a standard rubric developed explicitly for curriculum mapping, we compare the content of UMMedWiki with a validated medical education terminology developed by Willet *et al.*[15] Utilizing this terminology allowed for an assessment of the distribution of general topics included in medical education, as well as areas of possible bias, omissions, or duplications within the wiki. If the wiki can be shown with future research to reliably reflect the curriculum, then our metrics can be used to map curricular bias.

An initial reference database of diseases that contained common synonyms was created and duplications were removed. Subsequently, consolidation was completed after data selection. The primary reference source for common diseases was the AHRQ, which provided data on 100 diseases ranked by *incidence*.[23] AHRQ defined their

diseases with ICD-9 codes, which facilitated the addition of synonyms to our disease database. Since we preferred the more nuanced descriptions provided by SNOMED, we used the standardized map to convert the ICD-9 concepts into SNOMED concepts.

During the editing and evaluation phase, the entirety of this data was checked by hand for mapping inconsistencies. Duplications were eliminated, as were concepts that shared a common stem (since querying "infarct" also returned "infarction"). In approximately 5% of the disease data, we felt our rubric could be improved by reducing the specificity of the disease concept. Manifestations of a disease were converted to the disease itself (e.g. "cellulitis of leg" became "cellulitis"). Queries matched strings with identical word order and without intervening words. Since concatenation was allowed, abbreviations were used only cautiously. All deviations from standardized mapping were noted in the record.

The rare disease reference utilized the 2010 OrphaNet epidemiologic data set.[24] Attempts to find a US specific rare disease reference standard were unsuccessful and consistent with a 2008 finding by a study group convened by the Office of Rare Diseases Research that "found no broad compilation of data on the *prevalence* of disease in the US." The committee decided that OrphaNet, a European agency that developed a list of diseases ranked by *prevalence*, "had much in common with the NIH list of rare conditions." The committee determined that the OrphaNet data could be used as a proxy for disease *prevalence* within the US.[25] Unfortunately, most of the OrphaNet data did not include ICD-9 codes, and so we devised a script to specifically match concepts to SNOMED.

4) Results

4.1) Development: the Origin Story of UMMedWiki

The Medical School administration provides minimal support for the wiki, including free web hosting and a $600 budget for a volunteer student committee called the Knowledge Co-op (K-Coop). This group originally organized itself as a paper-based note sharing system. For every lecture, the group designated a particular student as the "note-taker" and copies of his notes were made and distributed to all participating cooperative students. When the wiki began in 2007 as a student-run project, it initially inherited the turn-taking structure of the note cooperative.

Although all students were invited to view and edit the wiki-based class notes, a core set of students initially alternated responsibility for editing the wiki as formally designated authors. The K-Coop continues to hold meetings for editors of the wiki, providing for minimal coordination. Currently, there is no formal turn-taking structure. Similar to other public wikis, no formal content review process is needed for submission or acceptance of an article. The editing policy of the UMMedWiki provides editing rights for any medical student, but recently students on the K-Coop voted to exclude faculty as editors, perhaps due to concerns that students would feel inhibited from making changes to faculty-authored content.

The majority of wiki articles summarize a particular lecture, but since no policy restricts content, the wiki has grown to include non-lecture material such as a library of normal lab values, a student petition to revise an unpopular class, and an outline of the student council budget. Since vast majority of source material for UMMedWiki is trustworthy lecture material, editors do not generally include citations for articles.

4.2) Usage: Who Visits and who Edits UMMedWIki?

As of March 9, 2011, Google Analytics reported 1,202,648 million total pageviews. This data includes the time course of pageviews, which lends reliability since expected seasonal patterns in usage can be seen following the academic calendar. Usage of the wiki peaks during early March, when students must choose their schedules for years 3 and 4. Prior to this scheduling, students appear to seek the student-authored course guides contained on UMMedWiki. While the overall number of pageviews peaked in 2008, the number of unique pageviews is more stable, peaking in 2009. Time spent on UMMedWiki is increasing and a new record length of use was recently established.

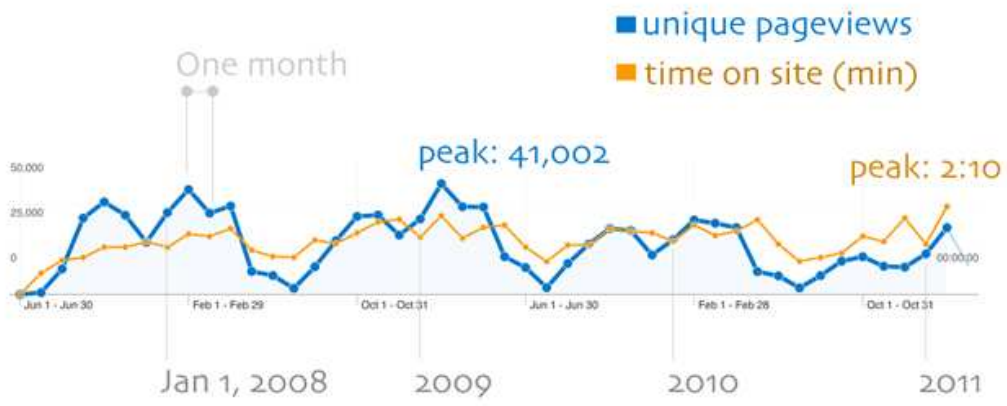


Figure 1

The user edit log showed that about 40 students have contributed a total of 28,000 edits to UMMedWiki; these 40 editors represent about 5% of the 781 potential users. Figure 3 shows data on the distribution of edits according to editor rank. One user, representing about 2.5% of the editor pool, contributed 25% of edits. Overall, however, 12% of editors contributed 50% of the edits, which is more egalitarian than reported by Chi *et al.* However, a line fitted to the data points of figure 3 using the equation $P(r)= 1/r^k$, shows that UMMedWiki's value for *k* equals 1.3, indicating egalitarianism within the range of Wikipedia as reported by a different set of researchers (k=0.65 to 1.65). Editor diversity peaked in 2008, and then dramatically declined to a steady state (graph on left below). The user edit log shows only 20 instances of article reversions—this suggests that less than .01 percent of edits repair overt damage.

Early phases of editing showed that the top ranked editor began work one month before the wiki was unveiled to the student body. This founding editor continued to contribute for an additional 18 months. Editors ranked five and six were not part of the initial surge of activity, although they became elite editors, contributing about 1,000 edits each into years 3 and 4 respectively.

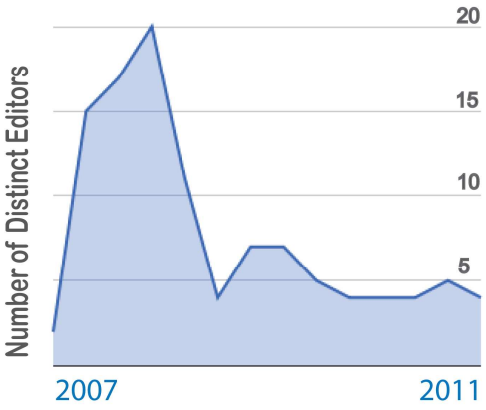


Figure 2

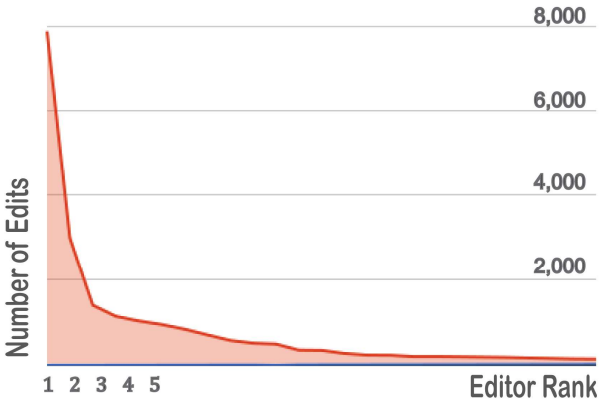


Figure 3

4.3) Content: UMMedWiki Articles

We found a total of 4,956 pages in UMMedWiki. This total includes pages that help facilitate wiki management including "talk" pages, meta-pages which are pages about the wiki itself, and pages with minimal content (text or images) called "stub" pages. These stub pages also include pages that redirect users to other pages. We arbitrarily defined "articles" as pages with a content length of greater than 300 characters, since the majority of these contained useful medical content (mostly lecture notes). If the page contained less than 300 characters, it was

much more likely to be a stub, as defined above. Only 1,571 pages had sufficient length to legitimately be termed "articles".

Characteristics of UMMedWiki articles help to contextualize its content relative to Wikipedia. The most salient characteristics define the meta-data used by algorithms to identify Featured Articles in Wikipedia. In table 1, we compare UMMedWiki's metadata to Wikipedia's average articles and Featured Articles. The best articles in UMMedWiki (75th percentile) compare favorably to Wikipedia's average articles, despite UMMedWiki's vastly smaller user base.

Metrics	Wikipedia's Exemplary[15]	Wikipedia's Average[15]	UMMedWiki Range	Avg.	25th	50th	75th
Number of unique editors	108	5	1-10		1	2	4.2
Total number of edits	257	8	2 - 223		2	6	18
Number of images	5	0	0 -5	2**			
Article length (chars)	24,708	1,344	300 - 37,000		705	1694	8250
Number of internal links	206	17	2-125		4	13	25

** Estimated from total number of uploaded images divided by number of articles. TABLE 1

4.4) Content: General Medical Topics

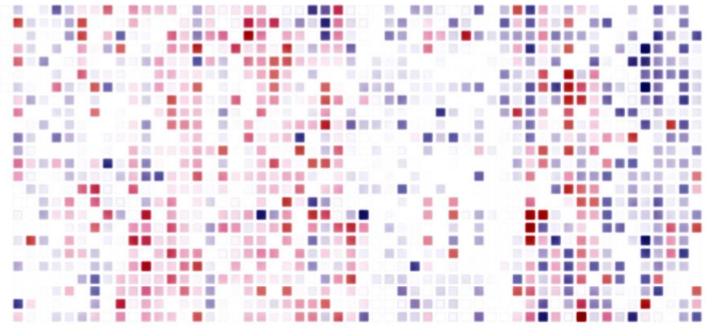
Surveying UMMedWiki with the TIME terminology showed that most concepts in the terminology could be mapped to the lecture notes or other miscellaneous content within UMMedWiki. Overall, 79% of the TIME concepts found at least one match in UMMedWiki, demonstrating that UMMedWiki contains broad coverage of concepts expected within a four year medical school curricula. Since we had also labeled the 531 TIME concepts as "disorders", we could gain an appreciation of how well UMMedWiki covered disorder concepts versus non-disorder concepts. The disorder concepts commonly known as diseases such as, "pneumonia", "diverticulitis", and "Noonan's Syndrome" while the non-disorder concepts include topics like "histology", "ethics", and "the healthcare system". We found that UMMedWiki mapped 89% of disorder concepts in TIME while only 70% of non-disorder concepts.

Each square represents a single TIME concept:

1 3
2 4

1.Most concepts are colored blue
2.except that disease concepts are red
3.Gaps exist where TIME failed to match.
4.The opacity varies by number of matches.

Visualization of 1,357 TIME concepts within UMMedWiki



Clusters: Although unintentional, some clustering is observed due to the sorting of child-parent relationships within the data file.

The square representing the concept "pain" is fully opaque because we found 280 articles in UMMedWiki that contained this matching TIME concept. Meanwhile, the concept "pleurisy" is faintly visible because we found this concept in only one article. Figure 4

4.5) Content: Survey of diseases

We found 13 relatively prevalent diseases in the OrphaNet list; these afflict about one in 3,300 people, yet had no reference in UMMedWiki. Of these, ten appeared to be clinically significant candidates for inclusion within the curriculum. We also found 19 candidates for over-representation. These diseases are cited more than once despite their *prevalence* of one in two million or less. These include, of note, Plummer Vinson may be as rare as one

per seven million, leading to doubts about its continued existence.[28] The candidates for under-representation were at least 600 times more prevalent than the candidates for over-representation. Table 2 and figure 4 summarize this data.

Disease Concept	Burden of Disease	# Articles with Citations
Chordoma, Fibrous Dysplasia of Bone, Tangier, Alexander and Castleman Diseases, Cutis Laxa. Gray Platelet, Li-Fraumeni, Muir-Torre, Plummer Vinson, and Bernard-Soulier Syndromes; Rhabdoid tumor.	Prevalence < .05 per 100,000	2-8
Diverticulitis, Spondylolisthesis, Non-Infectious Gastritis, Spinal Stenosis, Uterine Fibroid, Sinus Node Dysfunction, AV Fistula, Small Bowel Obstruction, Disc Degeneration, Premature Membrane Rupture, Schizoaffective Disorder.	Incidence ranges from 220,000 to 64,000 cases per year.	2-7
Congenital Absence of Vas Deferens, Hyperplastic Polyposis Syndrome, Obesity due to Melanocortin-4 Receptor Deficiency, Familial Melanoma, Congenital Thyroxine Binding Deficiency, Chromosome Y deletion, Familial Parkinson's Disease, Young Adult Onset of Parkinson's, Radiation Proctitis.	Prevalence of 50 per 100,000	0

Table 2

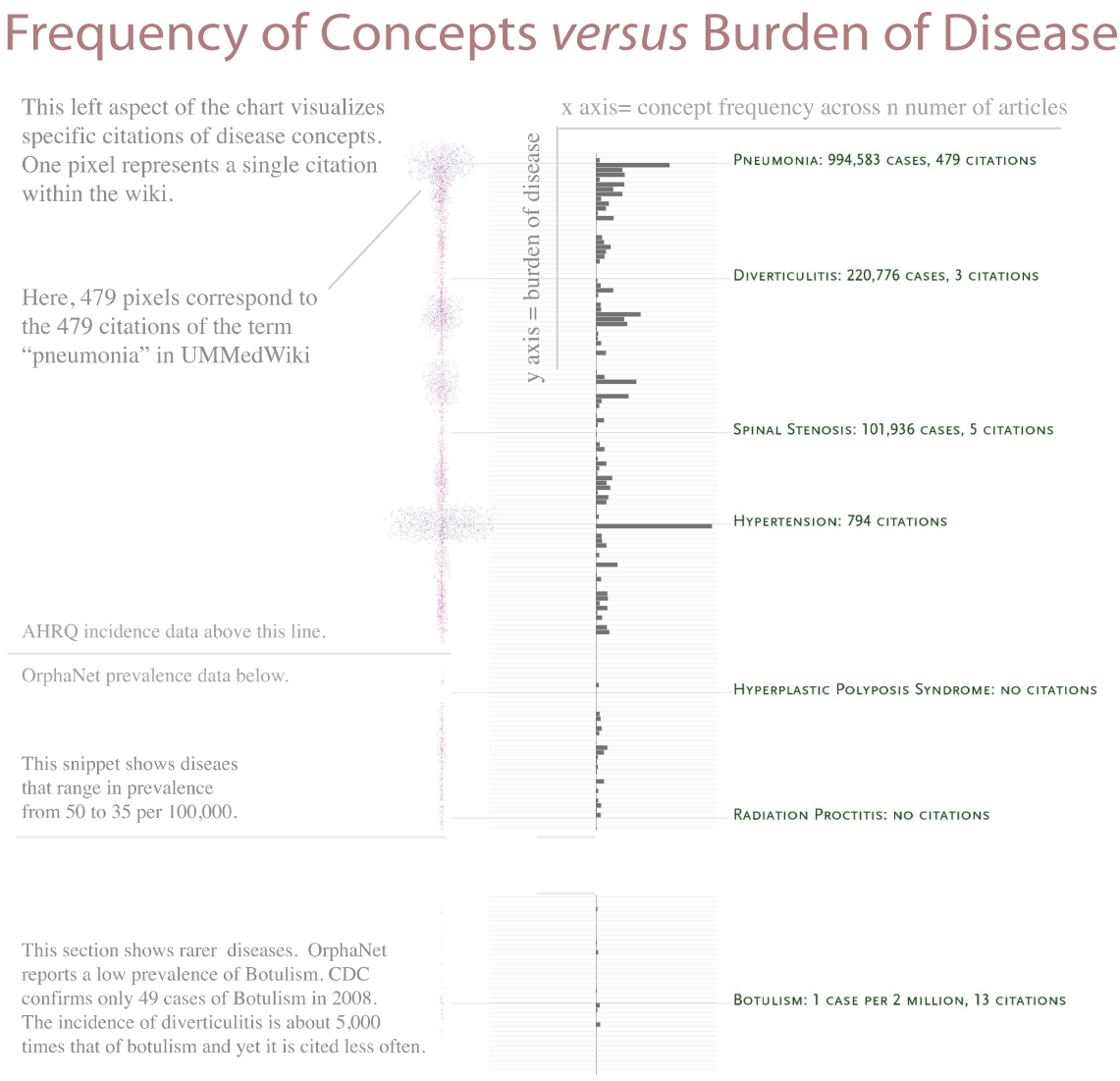


Figure 5

5) Discussion

UMMedWiki shows remarkable viewership-- 1.2 million pageviews-- despite minimal administrative funding. The size and egalitarianism of UMMedWiki's k-group, appears roughly correlated to Wikipedia. With regards to size, more information is needed here since we compare UMMedWiki's potential viewers (those given access to the site) to Wikipedia's actual viewers. Despite the success in terms of viewership and content, only 5% of potential viewers have made even small contributions, suggesting only larger schools will have the critical mass for an all-volunteer wiki. The most strikingly divergent dynamic within UMMedwiki is the rarity of vandalism. Coordination by the K-Coop, may have reduced conflict. Quality indicators-- such as article length, number of revisions, and editor diversity-- show similarity to an average article in Wikipedia, despite UMMedWiki's vastly smaller scale.

Early in its development, UMMedWiki benefited from an influential "founder" who operated at the inception of the wiki and contributed 25% of the edits. (This single editor constituted 2.5% of the editor pool.) The wiki now appears self-sustaining, albeit at a lower level of participation: more than 40 editors have contributed to the wiki over four years and some of the most prolific editors became involved within the last year.

Although common experience at the medical school suggests that the wiki is a comprehensive, accurate reflection of the curriculum, we have not provided direct evidence in this paper. However, TIME terminology mapped 89% of its disease concepts to the wiki, suggesting a fairly thorough coverage of expected concepts. TIME was designed for a four-year curriculum, while the wiki primarily covers the lecture-based initial years of medical school. Some discrepancy between disease burden and instruction was expected considering Denny's study that found a 40-fold bias favoring discussion of prions in lecture material compared to clinical notes. [21]

The long tail of the disease graph—which includes extremely rare diseases—appears to represent a cataloguing of diseases. A more explicit catalogue, perhaps coordinated by survey tools embedded in the wiki could help reduce omissions and areas of over-representation.

Our techniques for examining usage and content have some limitations. First, we use Number-of-Edits as a proxy for "value" because it is easier to calculate. As Preidhorsky et al. noted, ranking editors by the persistence and popularity of their contributions is better fit for the notion of user-created "value" than the more straightforward value, Number-of-Edits. Second, we rely on word counts instead of the more sophisticated use of natural language processing demonstrated by Denny. [21] This proxy interferes with mapping for a limited number of terms, such as "aspiration pneumonia" which may be extensively described without the immediate juxtaposition of "aspiration" and "pneumonia." Word counts that include synonyms probably suffice for many concepts, such as "Plummer Vinson" which are likely to be named specifically. While simple, our approach is transparent. Finally, our AHRQ hospitalizations data is somewhat unrefined since it does not aggregate codes describing different manifestations of the same disease. Although AHRQ provides a list of frequent hospitalizations based on aggregated codes, their approach produced concepts that appeared too general for our use.

In conclusion, the experience garnered through UMMedWiki shows how a self-sustaining wiki might be reproducible: an initial, prolific set of editors transfers content to the wiki; a turn-taking structure may promote editor diversity helping refine the wiki content. A student organization such as the Knowledge Co-op can coordinate the vision for the wiki, reducing conflict. We also show how an educational wiki might incorporate curriculum mapping to enhance its value. The student wiki may fill a unique educational niche providing a computable, comprehensive, and focused compendium on the core content of medical education. While it has clear student benefit, faculty might also find use by extracting reliable data for curriculum coordination and quality assurance. This might be enabled by allowing for faculty participation as nominators of "Featured Articles", thus providing a guide toward community quality standards.

Bibliography

1. E. Chi, "Long tail of user participation in Wikipedia," Retrieved March 30 (2007): 2009.
2. Jonathan L Streeter, Michael T Lu, and Frank J Rybicki, "Informatics in radiology: RadiologyWiki.org: the free radiology resource that anyone can edit," Radiographics: A Review Publication of the Radiological Society of North America, Inc 27, no. 4 (August 2007): 1193-1200.

3. M. Waldrop, “Big data: wikiomics,” Nature 455, no. 7209 (2008): 22–25.
4. Mitch Waldrop, “Big data: Wikiomics,” September 3, 2008, http://www.nature.com/news/2008/080903/full/455022a/table/1.html.
5. Jonathan L Streeter, Michael T Lu, and Frank J Rybicki, “Informatics in radiology: RadiologyWiki.org: the free radiology resource that anyone can edit,” Radiographics: A Review Publication of the Radiological Society of North America, Inc 27, no. 4 (August 2007): 1193-1200.
6. K. A Clauson et al., “Scope, completeness, and accuracy of drug information in Wikipedia,” The Annals of pharmacotherapy 42, no. 12 (2008): 1814.
7. Udo Altmann, “Representation of medical informatics in the wikipedia and its perspectives,” Studies in Health Technology and Informatics 116 (2005): 755-760.
8. Albert A Salas et al., “CurrMIT: a tool for managing medical school curricula,” Academic Medicine: Journal of the Association of American Medical Colleges 78, no. 3 (March 2003): 275-279.
9. Joshua C Denny et al., “"Where do we teach what?" Finding broad concepts in the medical school curriculum,” Journal of General Internal Medicine 20, no. 10 (October 2005): 943-946.
10. M. Y Lee et al., “Tufts Health Sciences Database: lessons, issues, and opportunities,” Academic Medicine 78, no. 3 (2003): 254.
11. J. Jacobs et al., “Implementing an online curriculum management database in a problem-based learning curriculum,” Academic Medicine 80, no. 9 (2005): 840.
12. D. R Graber et al., “Academic deans' views on curriculum content in medical schools,” Academic Medicine 72, no. 10 (1997): 901.
13. Leonard A Jason et al., “Frequency and content analysis of chronic fatigue syndrome in medical text books,” Australian Journal of Primary Health 16, no. 2 (2010): 174-178.
14. B. Stvilia et al., “Assessing information quality of a community-based encyclopedia,” in Proceedings of the International Conference on Information Quality, 2005, 442–454.
15. T. G Willett et al., “It's about TIME: a general-purpose taxonomy of subjects in medical education,” Medical education 42, no. 4 (2008): 432–438.
16. J. Wales, Wikipedia, emergence, and the wisdom of crowds, 2005.
17. R. B Almeida et al., “On the evolution of wikipedia,” in *Int. Conf. on Weblogs and Social Media*, 2007.
18. R. Priedhorsky et al., “Creating, destroying, and restoring value in Wikipedia,” in Proceedings of the 2007 international ACM conference on Supporting group work, 2007, 259–268.
19. P. E Dans and M. R Kerr, “Gerontology and geriatrics in medical education,” New England Journal of Medicine 300, no. 5 (1979): 228–232.
20. Janet B Henrich et al., “What do medical schools teach about women's health and gender differences?,” Academic Medicine: Journal of the Association of American Medical Colleges 81, no. 5 (May 2006): 476-482.
21. J. C Denny et al., “Comparing Content Coverage in Medical Curriculum to Trainee-Authored Clinical Notes,” in AMIA Annual Symposium Proceedings, vol. 2010, 2010, 157.
22. A. Haig et al., “METRO—the creation of a taxonomy for medical education,” Health Information & Libraries Journal 21, no. 4 (2004): 211–219.
23. “HCUPnet. Healthcare Cost and Utilization Project (HCUP)” (Agency for Healthcare Research and Quality, Rockville, MD., 2008), http://hcupnet.ahrq.gov.
24. “Orphanet,” http://www.orpha.net/consor/cgi-bin/index.php.
25. Marilyn J. Field and Thomas F. Boat, Editors; Committee on Accelerating Rare Diseases Research and Orphan Product Development; Institute of Medicine, Rare Diseases and Orphan Products: Accelerating Research and Development (Washington, D.C.: The National Academies Press, 2011).
26. L. S Buriol et al., “Temporal analysis of the wikigraph,” in Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence, 2006, 45–51.
27. T S Chen and P S Chen, “Rise and fall of the Plummer-Vinson syndrome,” Journal of Gastroenterology and Hepatology 9, no. 6 (December 1994): 654-658.
28. J. Giles, “Internet encyclopaedias go head to head,” *Nature* 438, no. 7070 (2005): 900–901.