

Matching Medical Websites to Medical Guidelines through Clinical Vocabularies in the Context of Website Quality Assessment

Dušan Rak

1.LF UK Praha, VŠE Praha



Problem

- **Information on Health Topics**
 - critical, sensitive, very important
 - extensive impact on health itself
- **Web:**
 - everything is on the Internet (same for the health topics)
 - uncontrolled creation
 - > huge amount of pages
 - > various target audience
 - > non-transparent responsibility
 - > various quality ! (even false info)
 - >>> **too many versions of truth !**



Progress in Medicine

- based on results of **scientific research**
- Evidence Based Medicine (**EBM**)
 - Techniques from Science, Engineering and Statistics
 - major method: Systematic review of published research studies
- hierarchy, **ranking of evidence**
 - strength of the freedom from various research biases
e.g.: US Preventive Services Task Force, National Health Service, [GRADE](#)
- Medical Guidelines (**MGL**) > **single version of truth !**

Reality

- **Individual Healthcare**

- Only some parts of HC are subject to scientific methods. EBM is just support for decision making (best prediction of treatment outputs)
- Many aspects depend on individual subjective factors (quality- and value-of-life judgments)
- Physicians need to put together their knowledge/experience, EBM best practice, patient (E)HR data and patient subjective inputs
- Final decision and responsibility will never be replaced by computer !
- Web resources have big influence on both patients and doctors !



What is Quality ?

- Quality: „value it delivers to the user“
- >> IQ is subjective
- Dimensions:

Intrinsic IQ (Accuracy, Objectivity, Believability...)

Contextual IQ (Relevancy, Value-Added, Timeliness, Completeness...)

Representational IQ (Interpretability, Ease of understanding, Concise representation...)

Accessibility



Quality of MGLs

- full version on the Internet
- regular updates
- EBM knowledge
- use of terminology
- well structured, electronic version GLIF model (can be used in EHR)
- GL metadata (autor, responsibility, validity..)
- NGC – differences, topic coverage, strength of evidence)

IQ on the Web

- often poor
 - >>>
- publication rules and guidelines for medicine on web
 - [quality criteria](#)
 - usually formal aspects only
 - no enforceability
- content evaluation:
 - **certification authorities** and **specialized catalogues**
 - manual expert work
 - no standard measures of content IQ

idea

- MGL can be used as a quality standard
- information content of text is represented by the terminology used
- comparison of terminology from MGLs vs. WS
- automatic determination of certain criteria
- support for expert decisionmaking

process

1. MGL

information: OK

used terminology: OK

--> “correct set of terminology” (standard)

2. Web

information: ?

terminology: ?

--> set of terms for each of documents

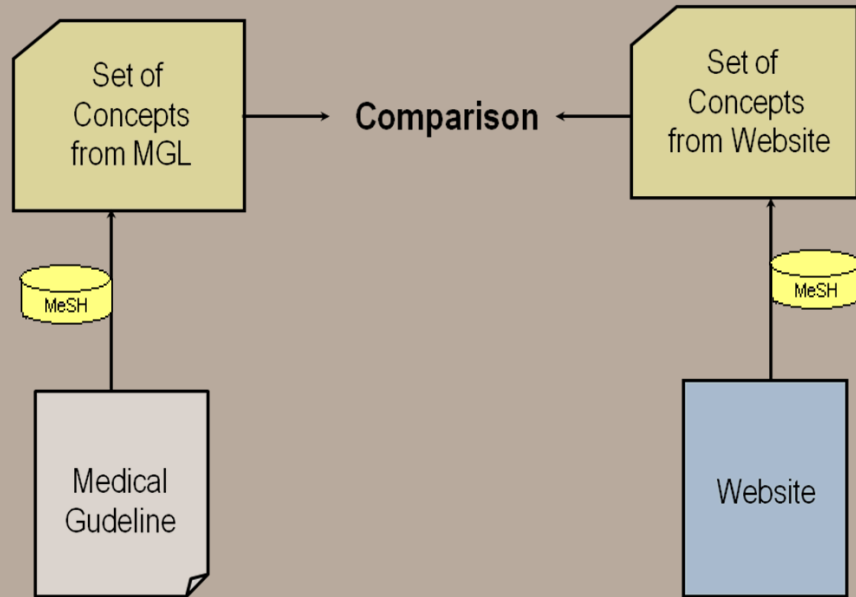
3. Porovnání obou výstupů

comparison

Content measures:

- use of appropriate terms
- scope of coverage of the main topic
- relevant/wrong treatment appears
- relevant/wrong drugs appear

idea



process I.

- model topic> “Screening for Lung Cancer”
- identification of guidelines> *3 found*
- web search + corpus creation > *top 100 from Google used*
- GL mapping to UMLS/MeSH: *concepts and synonyms for topic*
- 1st mapping of WS by UMLS/MeSH terms: *annotated documents*
- manual check of few annotations >>> *new synonyms*
- > adjusted hierarchy of concept-synonyms
- ...

process I.

MeSH, UMLS, ICD, SNOMED CT	Lung Cancer Screening	Tomography, Spiral Computed	Radiography, Thoracic	ology	nchoscopy	ath Tests	Lung Neoplasms	
concept	Lung Cancer	CT Scan	chest X-ray	MeSH: D036542 UMLS: HCPCS: S8092 Electron beam computed tomography (also known as Ultrafast CT, Cine CT) ICD: SNOMED-CT: synonyms: CAT Scan, Spiral Computed Tomography, Spiral Computer-Assisted Tomography, Spiral Computerized Tomography, Spiral CT Scan, Spiral Helical Computed Tomography Helical CT Spiral Computed Tomography Spiral CT				lung cancer
	Screening Lung Cancer Screening Screening for Lung Cancer	scan CT -- CAT -- computed tomography low dose -- spiral -- helical -- LDCT	X-ray radiography CXR chest --					Carcinoma tumor

process II.

- ...
- mapping of all WS and MGLs to adjusted synonyms
- > sets of concepts/synonyms for each of texts
- 4 MGL aggregates tested
- cross similarity between 3 MGLs and 4 aggregates
- ...

	similarity (terms)						
document	$a1 \cap a2 \cap a0$	$a1 \cup a2 \cup a0$	$nsum(a1, a2, a0)$	$sum(a1, a2, a0)$	a0	a1	a2
$a1 \cap a2 \cap a0$	1,00	0,66	0,70	0,71	0,73	0,63	0,69
$a1 \cup a2 \cup a0$	0,66	1,00	0,58	0,61	0,64	0,44	0,64
$nsum(a1, a2, a0)$	0,70	0,58	1,00	1,00	0,99	0,95	0,95
$sum(a1, a2, a0)$	0,71	0,61	1,00	1,00	1,00	0,92	0,97
a0	0,73	0,64	0,99	1,00	1,00	0,89	0,98
a1	0,63	0,44	0,95	0,92	0,89	1,00	0,81
a2	0,69	0,64	0,95	0,97	0,98	0,81	1,00

process III.

- ...
- similarity of WS to benchmarks
 1. on concept level
 2. on synonym/terms level
 3. on distinct terms level
- additional WS target audience categories
- similarity for each category

type	description	count
m	WS for medical professionals	23
p	WS for patients	21
ch	WS for children	0
g	general, news, other	28
mo	scientific papers (full texts or at least abstract)	23
mo/x	scientific papers (restricted access, usually title only)	5

results

- method tested on single carefully selected topic
 - > some manual steps unavoidable
- 4 aggregation methods were evaluated
 - > “sum” and “nsum” appeared the best
- relation between similarity and target audience identified
 - > dependence showed
 - > WS categorisation and evaluation of similarity per category added to process workflow
- levels of similarity compared between analyses on concept, term and distinct term level
 - > concept level: evaluates the topic and scope match
 - > term level: evaluates terminology match

manual steps in process

- model topic
- identification of guidelines
- web search + corpus creation
- GL mapping to UMLS/MeSH – concept identification
- 1st mapping of WS by UMLS/MeSH terms: *annotated documents*
- manual check of few annotations >>> new synonyms
- > adjusted hierarchy of concept-synonyms
- mapping of all WS and MGLs to adjusted synonyms
- > sets of concepts/synonyms for each of texts
- 4 MGL aggregates tested
- cross similarity between 3 MGLs and 4 aggregates
- similarity of WS to benchmarks
- additional WS target audience categories
- similarity for each category

similarity per target audience

category / document	similarity (terms)							similarity (concepts)							similarity (distinct terms)						
	$a1 \cap a2 \cap a0$	$a1 \cup a2 \cup a0$	nsum (a1,a2,a0)	sum (a1,a2,a0)	a0	a1	a2	$a1 \cap a2 \cap a0$	$a1 \cup a2 \cup a0$	nsum (a1,a2,a0)	sum (a1,a2,a0)	a0	a1	a2	$a1 \cap a2 \cap a0$	$a1 \cup a2 \cup a0$	nsum (a1,a2,a0)	sum (a1,a2,a0)	a0	a1	a2
g	0,53	0,43	0,83	0,82	0,81	0,78	0,80	0,75	0,75	-	0,75	0,75	0,75	0,75	0,50	0,53	-	0,56	0,55	0,47	0,53
m	0,59	0,50	0,86	0,87	0,87	0,78	0,87	0,75	0,77	-	0,77	0,77	0,75	0,77	0,51	0,63	-	0,64	0,64	0,51	0,61
mo	0,60	0,53	0,88	0,89	0,89	0,80	0,88	0,72	0,79	-	0,79	0,79	0,72	0,79	0,52	0,67	-	0,67	0,65	0,52	0,66
ch	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
mo/x	0,57	0,48	0,85	0,85	0,86	0,76	0,85	0,64	0,73	-	0,73	0,73	0,64	0,73	0,50	0,53	-	0,57	0,54	0,47	0,56
p	0,56	0,45	0,84	0,83	0,82	0,79	0,81	0,75	0,73	-	0,73	0,73	0,75	0,73	0,52	0,57	-	0,60	0,59	0,49	0,55
all	0,57	0,48	0,85	0,85	0,85	0,79	0,84	0,74	0,76	-	0,76	0,76	0,74	0,76	0,51	0,59	-	0,61	0,60	0,49	0,58

Questions, discussion ?