

Indexing with semantic components improve information retrieval in domain-specific web portal

Marianne Lykke Nielsen
Information Interaction and Information Architecture
Royal School of Library and Information Science



NKOS Workshop 2008
At ECDL, Århus, Denmark
September 19, 2008

Agenda



- Indexing with semantic components
- Indexing and searching problems in domain-specific IR systems
- User evaluation: indexing and retrieval
- Results
- Summing up

Searching problems



- Domain-specific IR systems often contain a large number of documents about a limited set of topics – **difficult to distinguish** between documents
- Domain experts have often very **specific information needs** related to specific work tasks
- **Situational relevance**, determined by contextual dimensions is often more important than topical relevance
- **Limited time** for searching

Indexing problems



- **Topics** – what topics should be indexed
- **Exhaustivity and specificity** - what level of specificity and completeness (exhaustivity)
- **Terminology** – what terms and terminology should be used to express the topics
- **Consistency** – how do we ensure consistency
- **Costs** – indexing takes time and demands domain knowledge and indexing skills and knowledge

Indexing with semantic components (SC)



- Marking up semantic sections of text (*semantic components*)
- Sections contain information about specific aspects or facets of the overall topic, e.g. diagnosis, treatment, referral, risk factors
- Method based on assumptions and research results that show:
 - Domain experts know document types with a certain domain. They know document structure, and use this knowledge when reading and using documents (Dillon, 1991; Orlikowaki & Yates, 1994; Bishop, 1999)
 - Content and structure in domain-specific documents correspond to aspects/facets of domain-specific information needs (Ely et al, 1999,2000; Price, Delcambre, Nielsen, 2006)
- Supplementary to basic indexing methods, manual as well as automatic

Research team



General medicine

Peter Vedsted

MD, Ph.D.

Research Unit General Medicine,
Århus University

Jens Rubak

MD

Praksis.dk, Region Midtjylland

Information and computer science

Lois Delcambre, Ph.D., Professor

Susan Price, MD, Ph.D. student
Computer Science Department
Portland State University, USA

Marianne Lykke Nielsen, Ph.D., Asso. professor
Information Interaction and Architecture
Royal School of Library and Information Science,
Denmark

sundhed.dk

Vibeke Luk

Information specialist
sundhed.dk

Frans la Cour

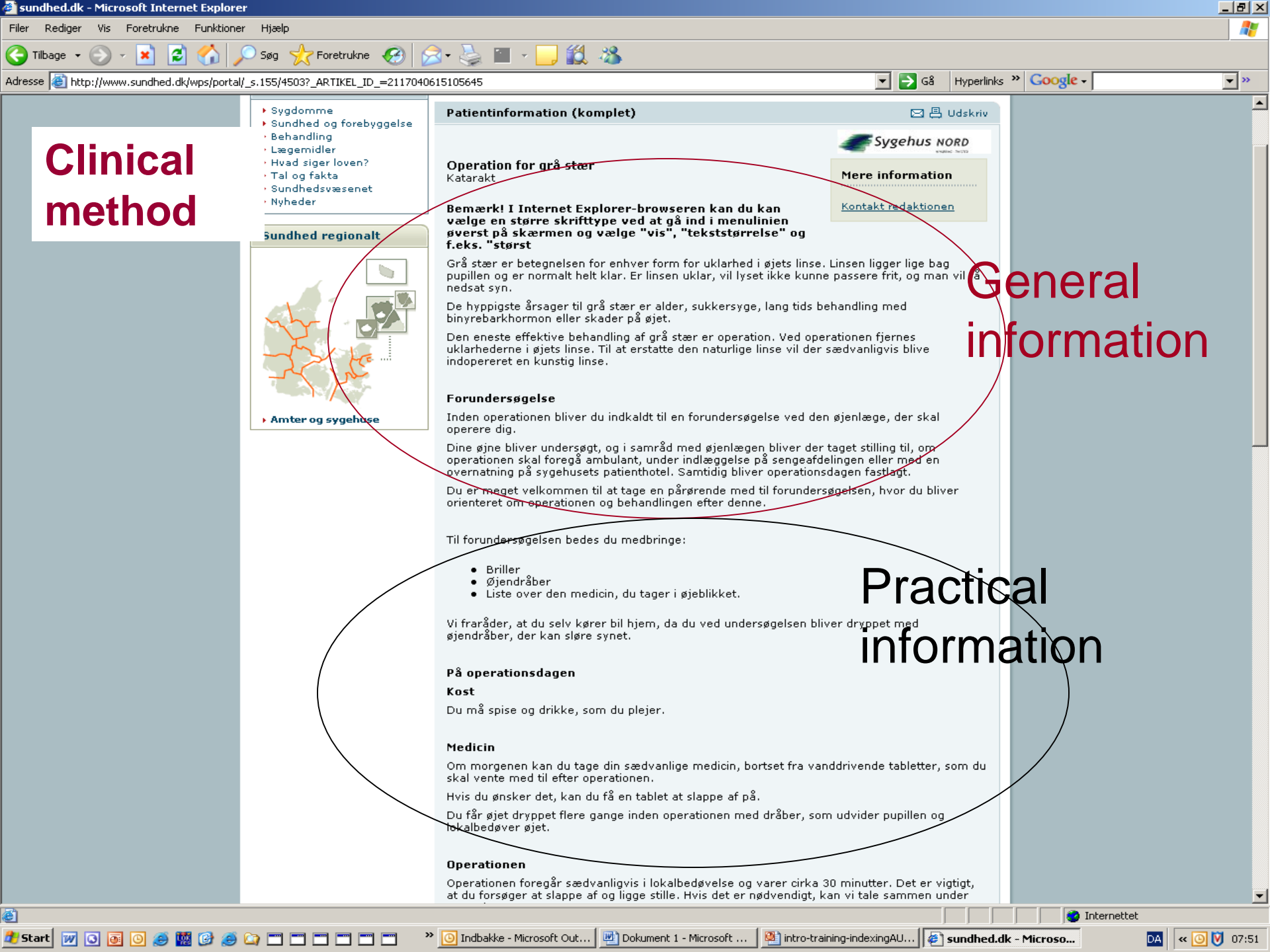
IT consultant
Autonomy

Supported by US National Science Foundation, sundhed.dk and Region Midtjylland

Case study



- sundhed.dk: national, Danish health portal
- Established in 2001, 25.000 documents
- Two main target groups: citizens and professionals in the health care sector
- Use of automatic as well as manual indexing methods:
 - ICPC
 - ICD-10
 - Citizens Thesaurus
- Large and varied group of indexers
 - 17 regions
 - Up to 250 indexers per region
- Research project focuses on family practitioners



Clinical
method

General
information

Practical
information

- Sygdomme
- Sundhed og forebyggelse
- Behandling
- Lægemidler
- Hvad siger loven?
- Tal og fakta
- Sundhedsvæsenet
- Nyheder



Patientinformation (komplet)

Udskriv

Sygehus NORD

Mere information

[Kontakt redaktionen](#)

Operation for grå stær Katarakt

Bemærk! I Internet Explorer-browseren kan du kan vælge en større skrifttype ved at gå ind i menulinien øverst på skærmen og vælge "vis", "tekststørrelse" og f.eks. "størst"

Grå stær er betegnelsen for enhver form for uklarhed i øjets linse. Linsen ligger lige bag pupillen og er normalt helt klar. Er linsen uklar, vil lyset ikke kunne passere frit, og man vil på nedsat syn.

De hyppigste årsager til grå stær er alder, sukkersyge, lang tids behandling med binyrebarkhormon eller skader på øjet.

Den eneste effektive behandling af grå stær er operation. Ved operationen fjernes uklarhederne i øjets linse. Til at erstatte den naturlige linse vil der sædvanligvis blive indopereret en kunstig linse.

Forundersøgelse

Inden operationen bliver du indkaldt til en forundersøgelse ved den øjenlæge, der skal operere dig.

Dine øjne bliver undersøgt, og i samråd med øjenlægen bliver der taget stilling til, om operationen skal foregå ambulant, under indlæggelse på sengeafdelingen eller med en overnatning på sygehusets patienthotel. Samtidig bliver operationsdagen fastlagt.

Du er meget velkommen til at tage en pårørende med til forundersøgelsen, hvor du bliver orienteret om operationen og behandlingen efter denne.

Til forundersøgelsen bedes du medbringe:

- Briller
- Øjendråber
- Liste over den medicin, du tager i øjeblikket.

Vi fraråder, at du selv kører bil hjem, da du ved undersøgelsen bliver dryppet med øjendråber, der kan sløre synet.

På operationsdagen

Kost

Du må spise og drikke, som du plejer.

Medicin

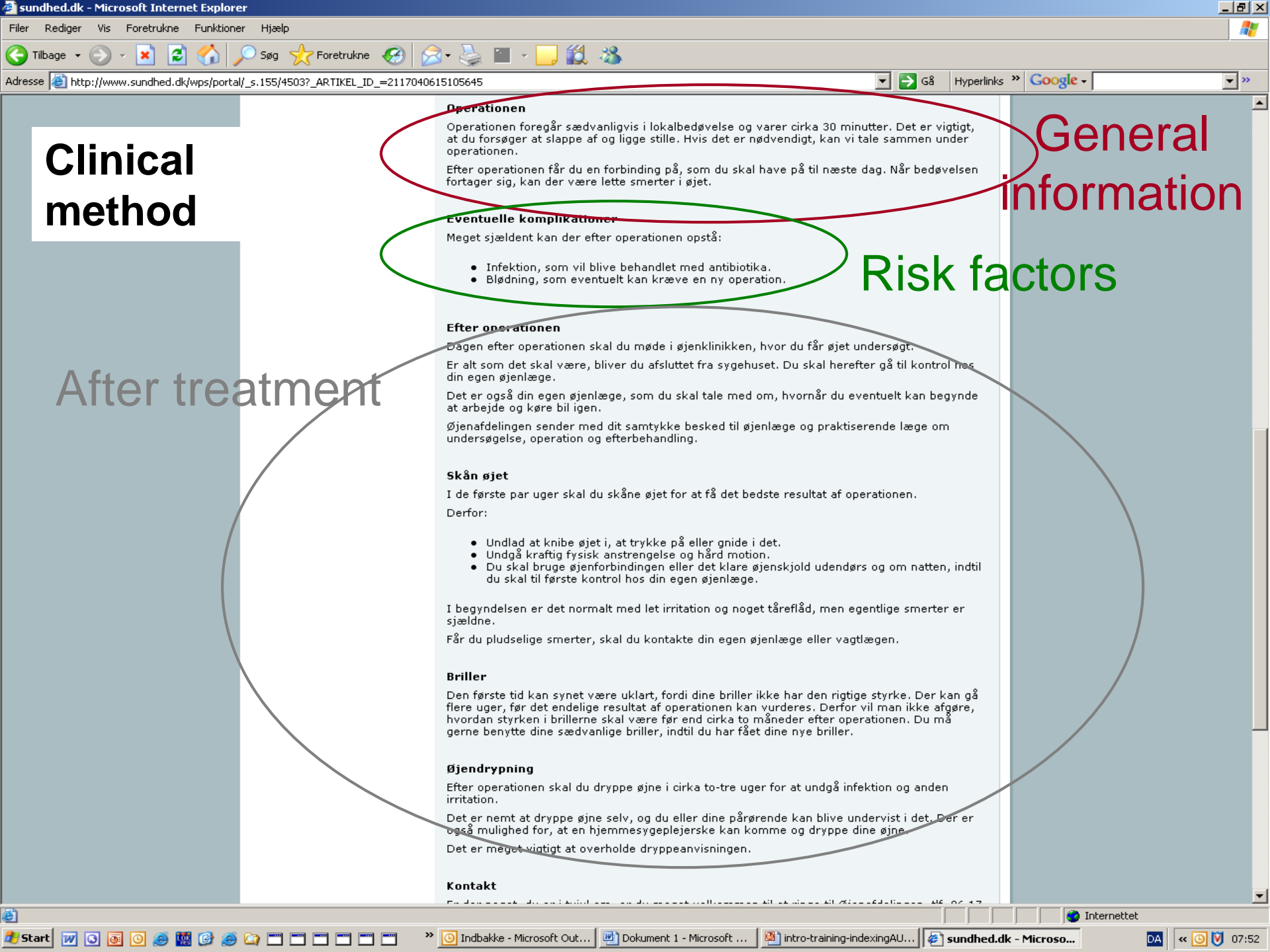
Om morgenen kan du tage din sædvanlige medicin, bortset fra vanddrivende tabletter, som du skal vente med til efter operationen.

Hvis du ønsker det, kan du få en tablet at slappe af på.

Du får øjet dryppet flere gange inden operationen med dråber, som udvider pupillen og lokalbedøver øjet.

Operationen

Operationen foregår sædvanligvis i lokalbedøvelse og varer cirka 30 minutter. Det er vigtigt, at du forsøger at slappe af og ligge stille. Hvis det er nødvendigt, kan vi tale sammen under



Clinical method

After treatment

General information

Risk factors

Operationen

Operationen foregår sædvanligvis i lokalbedøvelse og varer cirka 30 minutter. Det er vigtigt, at du forsøger at slappe af og ligge stille. Hvis det er nødvendigt, kan vi tale sammen under operationen.

Efter operationen får du en forbinding på, som du skal have på til næste dag. Når bedøvelsen fortager sig, kan der være lette smerter i øjet.

Eventuelle komplikationer

Meget sjældent kan der efter operationen opstå:

- Infektion, som vil blive behandlet med antibiotika.
- Blødning, som eventuelt kan kræve en ny operation.

Efter operationen

Dagen efter operationen skal du møde i øjenklinikken, hvor du får øjet undersøgt.

Er alt som det skal være, bliver du afsluttet fra sygehuset. Du skal herefter gå til kontrol hos din egen øjenlæge.

Det er også din egen øjenlæge, som du skal tale med om, hvornår du eventuelt kan begynde at arbejde og køre bil igen.

Øjenafdelingen sender med dit samtykke besked til øjenlæge og praktiserende læge om undersøgelse, operation og efterbehandling.

Skån øjet

I de første par uger skal du skåne øjet for at få det bedste resultat af operationen.

Derfor:

- Undlad at knibe øjet i, at trykke på eller gnide i det.
- Undgå kraftig fysisk anstrengelse og hård motion.
- Du skal bruge øjenforbindingen eller det klare øjenskold udendørs og om natten, indtil du skal til første kontrol hos din egen øjenlæge.

I begyndelsen er det normalt med let irritation og noget tåreflåd, men egentlige smerter er sjældne.

Får du pludselige smerter, skal du kontakte din egen øjenlæge eller vagtlægen.

Briller

Den første tid kan synet være uklart, fordi dine briller ikke har den rigtige styrke. Der kan gå flere uger, før det endelige resultat af operationen kan vurderes. Derfor vil man ikke afgøre, hvordan styrken i brillerne skal være før end cirka to måneder efter operationen. Du må gerne benytte dine sædvanlige briller, indtil du har fået dine nye briller.

Øjendrypning

Efter operationen skal du dryppe øjne i cirka to-tre uger for at undgå infektion og anden irritation.

Det er nemt at dryppe øjne selv, og du eller dine pårørende kan blive undervist i det. Der er også mulighed for, at en hjemmesygeplejerske kan komme og dryppe dine øjne.

Det er meget vigtigt at overholde dryppeanvisningen.

Kontakt

Følgende telefonnumre kan du kontakte, hvis du har spørgsmål til operationen eller efterbehandling.

Semantic components (SC) in sundhed.dk



- In sundhed.dk we have identified 6 document types:
 - *Clinical problem*
 - *Clinical method*
 - *Services*
 - *Drug product*
 - *Clinical unit*
 - *Note*
- Each document type has its own set of semantic components, for example *Clinical method* contains 6 semantic components:
 - *General information*
 - *Practical information*
 - *Referral*
 - *After treatment*
 - *Risk factors*
 - *Expected result*

Document

(1) Enter Sundhed.dk document id

(2) Press here to get document

Kronisk Obstruktiv Lungelidelse (State of the Art)

Indhold:

- ◆ [Prolog](#)
- ◆ [1.0 Kronisk obstruktiv lungesygdom](#)
- ◆ [2.0 Lidelsens omfang og omkostninger](#)
- ◆ [3.0 Diagnostik](#)
 - [3.1 Lungefunktionsundersøgelse](#)
 - [3.2 Røntgenundersøgelse](#)
- ◆ [4.0 Klinisk vurdering](#)
 - [4.1 Individuel vurdering af den gavnlige virkning af symptomatisk behandling](#)
 - [4.2 Flowchart](#)
- ◆ [5.0 Behandling](#)
 - [5.1 Forebyggende behandling inkl. risikofaktorer](#)
 - [5.2 Non-farmakologisk behandling](#)
 - [5.3 Farmakologisk behandling](#)
 - [5.4 Rehabilitering af KOL-patienter](#)
- ◆ [6.0 Rygning](#)
- ◆ [7.0 KOL og obstruktiv søvnapnø](#)
- ◆ [Litteraturreferencer](#)
- ◆ [Forfattere](#)
- ◆ [Bilag 1. Patientoplysningsmateriale](#)
- ◆ [Bilag 2. Inhalationssystemer](#)
- ◆ [Bilag 3. Valg af inhalator](#)

Prolog [\(til top\)](#)

>Under overskriften fremme af god klinisk praksis har en række arbejdsgrupper beskrevet "state of the art" på flere områder. Arbejdet foregår i regi af Kvalitetsudviklingsudvalget vedrørende almen praksis i Århus Amt.

Dokumentet er et tilbud til de praktiserende læger og deres efteruddannelsesgrupper, som kan bruge dokumentet til efteruddannelse og kvalitetsudvikling.

Semantic Components

Title

(3) Enter your indexer id

(4) Choose the type for the main topic of the document:

- Klinisk enhed
- Klinisk problem**
- Klinisk metode
- Services
- Opslag
- Lægemidler

(5) Click here to start indexing

Document

(1) Enter Sundhed.dk document id Aarhus_Amt1014040716133826.html

(2) Press here to get document

70-74 år	75-79 år	80-84 år	85-89 år
204	429		27
I alt	5265	7287	549

Kilde: Sygesikringens data fra afregning med apotekerne

I 2003 har 86 % af amtets alment praktiserende læger målt spirometri. Til sammenligning målte 59 % af lægerne spirometri 1997.

Sygesikringsudgiften til midler mod astma/KOL til de +50 årige var i 2003 62,6 mill. kr. svarende til et forbrug på 8,9 mill. Definerede Døgn Doser (DDD).

Også i hospitalssystemet bruges som nævnt mange ressourcer på behandling af kroniske lungepatienter.

*Kroniske lungepatienter *) i 2003 fordelt på sygehusene i Århus Amt*

	Aktionsdiagnose	Bidiagnose	Sengedage	"Omkostninger"
Alle sygehuse	1727	2903	28.836	122,4 mill. kr.

Note *) Antal patienter, hvor der er stillet diagnosen DJ43 - udvidelse af lunger eller DJ44 - kronisk obstruktiv lungesygdom, anden. Omkostninger beregnet på grundlag af DRG-værdier.

>

>3.0 Diagnostik (til top)

KOL defineret som $FEV1/FVC < 0,7$ uden tegn på astma,
>kan kun diagnosticeres ved
>spirometri

>Hos nogle patienter med KOL kan der ses en bedring i lungefunktionen efter bronkieudvidende midler. Dette benævnes KOL med reversibilitet. Tilstedeværelsen af reversibilitet er dog ikke afgørende for behandling med bronkiedilatatorer eller inhalationssteroid, idet der sker mange andre lungefunktionsforandringer som kan nedsætte respirationsarbejdet. Det vigtigste er derfor den kliniske respons på behandling.

Størrelsen af FEV1 er et mål for obstruktionens sværhedsgrad, og gentagne målinger af FEV1 er anvendeligt til at monitorere obstruktionens udvikling. FEV1 er prædikator for overlevelsen. Ved FEV1 på 1,0 liter er der en median overlevelse på 5 år, og ved FEV1 på 0,5 liter er medianoverlevelsen på mindre end 2 år. Prognosen er dog mest afhængig af

Semantic Components

(6) Click here when you have finished indexing this document

Generel information

Erase this instance of Generel information

1.0 Kronisk obstruktiv lungesygdom (til top)

>Kronisk obstruktiv lungesygdom (KOL) er en tilstand med patologiske forandringer i luftvejene og lungeparenkymet, som medfører nedsat maksimal luftstrømningshastighed. Kronisk bronkitis, emfysem og blandingstilstande heraf resulterer i en progredierende nedsat lungefunktion og deraf følgende vejrtrækningsbesvær. Symptomerne er åndenød ved anstrengelse, hoste og opspyt.

>Da det drejer sig om kroniske, irreversible lungeskader, er lungefunktionsnedsættelsen ikke fuldt reversibel, sådan som det i almindelighed er tilfældet ved astma. Dog kan mangeårig astma medføre irreversible skader således, at man ikke med spirometer alene kan skelne astma fra KOL.

Diagnosticering

Erase this instance of Diagnosticering

Størrelsen af FEV1 er et mål for obstruktionens sværhedsgrad, og gentagne målinger af FEV1 er anvendeligt til at monitorere obstruktionens udvikling. FEV1 er prædikator for overlevelsen. Ved FEV1 på 1,0 liter er der en median overlevelse på 5 år, og ved FEV1 på 0,5 liter er medianoverlevelsen på mindre end 2 år. Prognosen er dog mest afhængig af andre forhold såsom den almindelige fysiske kondition, den kardiovaskulære tilstand og ernæringstilstanden.

KOL opdeles i 4 sværhedsgrader. Sværhedsgraden er en definition af, hvor alvorlig sygdommen er hos den enkelte, og vurderes på grundlag af symptomhyppighed, påvirkning af daglige funktioner, natlige symptomer, lungefunktion og forbrug af akut medicin. Behandlingen tilrettelægges ud fra sværhedsgraden.

Henviisning

Erase this instance of Henviisning

Behandling

Erase this instance of Behandling

Search:

astma

in addition to your main search term ...

Afgræns søgning:

Informationstype: Alle typer

Regionalt indhold: Alt regionalt indhold

... enter a term to search within a semantic component

Enter terms to search within semantic

Klinisk problem

- Generel information
- Diagnosticering
- Henvisning
- Behandling

Services

- Generel information
- Praktisk information
- Henvisning

Klinisk metode

- Generel information
- Praktisk information
- Henvisning
- Efterbehandling
- Risici
- Forventet resultat

Klinisk enhed

- Funktion og speciale
- Praktisk information
- Henvisning
- Personale og organisation

Lægemidler

- Generel information
- Praktisk information
- Målgruppe
- Effekt
- Bevirkning, Interaktioner og kontraindikationer

Opslag

- Generel information
- Praktisk information
- Kvalifikation

Search:

Afgræns søgning:

Informationstype:

Regionalt indhold:

Enter terms to search within semantic componen

Klinisk problem

Generel information

Diagnosticering

Henvisning

Behandling

Services

Generel information

*



Enter an asterisk to find all (and only) the **Clinical problem** documents about asthma that contain the **Referral** semantic component

Klinisk m

Generel

Praktisk

Henvisn

Efterbeh

Risici

Forvent

Klinisk e

Funktion

Praktisk

Henvisn

Persona

Lægemid

Generel

Praktisk

Målgrup

Effekt

Bevirkni

Opslag

User evaluation - searching



- **Comparison** between standard, control retrieval system (System 1) and experimental retrieval system with semantic components (system 2)
- **Test persons:** 30 family practitioners with experience with sundhed.dk and online information retrieval
- **Training sessions:** introduction to semantic components and the two retrieval systems
- **Search jobs:** 4 controlled, simulated search scenarios
 - 2 search jobs are carried out in System 1
 - 2 search jobs are carried out in System 2
- **Random distribution** of search jobs and retrieval systems
- **Data collection:**
 - Searching behavior, graded relevance assessments (user and system relevance), time used, ease of use, and user satisfaction)

Search scenarios



Search scenario C

Anna is childless. She had two abortions (un-intentional). She is now ready to get pregnant again. There is something about folate. Should she have it, and what doses?

Search task: Find documents that help you to decide, whether Anna should take folate, and how big a doses.

Test persons

Test persons (n = 30)				
	Experience Internet search engines (year)	Experience sundhed.dk (year)	Experience Information retrieval (1 Much – 5 None)	Experience Medical professional (year)
Mean	7,2	2,4	2,4	21,4
Maximum	12,5	5	4	34,2
Minimum	1	0,6	1	6,3
Std	2,82	1,43	0,93	7,59

Search performance

**MAP for Best query per session
(Mean \pm SE)**

	Search scenario A	Search scenario B	Search scenario C	Search scenario D	Total
System 1	0.28 \pm 0.03	0.53 \pm 0.11	0.21 \pm 0.05	0.19 \pm 0.03	0.31 \pm 0.03
System 2	0.56 \pm 0.06	0.58 \pm 0.11	0.26 \pm 0.03	0.27 \pm 0.06	0.42 \pm 0.04

**nDCG for Best query per session
(Mean \pm SE)**

	Search scenario A	Search scenario B	Search scenario C	Search scenario D	Total
System 1	0.41 \pm 0.03	0.57 \pm 0.08	0.38 \pm 0.07	0.34 \pm 0.03	0.43 \pm 0.03
System 2	0.60 \pm 0.06	0.60 \pm 0.08	0.50 \pm 0.04	0.46 \pm 0.07	0.54 \pm 0.03

Ease of use

Ease of use, measured by queries and time
(n = 120)

	No of queries per session		Time used (minutes) per session	
	System 1 (n = 60)	System 2 (n = 60)	System 1 (n = 60)	System 2 (n = 60)
Mean	2,55	3,2	5,8	7,2
Minimum	1	1	1,9	2,2
Maximum	8	11	16,8	16,6
Std	1,908	2,386	-	-

Ease of use

Ease of use (1 Good – 5 Bad)						
	How easy to formulate	How satisfied with result	How useful to specify with Information type	How useful to specify with Region	How useful to specify with SC	How useful to search with SC
System 1	2,02	2,12	1,90	1,54	-	-
System 2	2,07	2,28	1,62	1,65	1,58	1,73

User satisfaction

- 73% of users indicate that they will use SC for more than 50% of their daily searches
- SC is useful for specific search tasks
 - SC is useful for complex, specific tasks – for tasks where you are “on new and unknown territory”
- SC is not intuitive, demands training
 - Users need training to use SC effectively
 - Labels are not intuitive
- Better functionality
 - More simple: fewer SC and no division into document types
 - Personalization
 - Direct access from interface to semantic component in the text

User evaluation - indexing



- Comparison of keyword indexing with indexing with semantic components
- Test persons: 16 Danish sundhed.dk indexers
- Training sessions: introduction to indexing with semantic components (SC)
- Indexing tasks: 12 sundhed.dk documents
 - 6 documents were indexed by semantic components (SC)
 - 6 documents were indexed by keywords: free, ICPC, ICD-10, domain-specific thesaurus (Borgertesaurus)
- Random distribution of indexing tasks and indexing methods
- Data collection:
 - Indexing data
 - Time used
 - User satisfaction

Evaluation metrics - indexing

Indexing method	Accuracy (by use of reference standard)	Consistency
Semantic components (SC)	Recall and Precision - Of characters in each component	Binary K - Agreement of marking up a semantic component - Over all characters in each component
Keywords	Recall and Precision - keywords	Binary K - Agreement of keyword assignment - Over all keywords, selected by one indexer or in reference standard Traditional consistency metrics

Akkurathed – semantisk komponenter

Document	Semantic component		Keywords	
	Recall macroaverage	Precision macroaverage	Recall macroaverage	Precision macroaverage
1	0.74 ± 0.37	0.89 ± 0.26	0.14 ± 0.33	0.74 ± 0.43
2	0.56 ± 0.33	0.61 ± 0.39	0.35 ± 0.47	0.74 ± 0.42
3	0.59 ± 0.45	0.72 ± 0.38	0.10 ± 0.23	0.72 ± 0.42
4	0.33 ± 0.29	0.72 ± 0.41	0.16 ± 0.35	0.70 ± 0.45
5	0.74 ± 0.39	0.68 ± 0.47	0.38 ± 0.47	0.85 ± 0.30
6	0.59 ± 0.13	0.81 ± 0.35	0.01 ± 0.04	0.88 ± 0.31
7	0.63 ± 0.39	0.79 ± 0.31	0.28 ± 0.36	0.62 ± 0.41
8	0.70 ± 0.31	0.93 ± 0.17	0.01 ± 0.02	0.61 ± 0.49
9	0.66 ± 0.33	0.76 ± 0.43	0.21 ± 0.39	0.79 ± 0.39
10	0.61 ± 0.35	0.75 ± 0.26	0.25 ± 0.42	0.79 ± 0.39
11	0.65 ± 0.43	0.86 ± 0.31	0.12 ± 0.27	0.80 ± 0.36
12	0.63 ± 0.48	0.83 ± 0.30	0.03 ± 0.08	0.85 ± 0.34

Summing up



- We developed the method *Indexing with semantic components* – to improve precision in retrieval
- We tested the retrieval effectiveness of the method in a case study with 30 family practitioners in sundhed.dk
 - Over all scenarios, the improvement in was 35.5% as measured by MAP and 25.6% by nDCG – compared with the standard, control system
- We tested indexing accuracy in an indexing test with 16 sundhed.dk indexers
 - Findings indicate that it is easier to obtain higher accuracy with semantic component indexing as regards recall
- Future research
 - Develop the method, e.g. simpler version, automation of up marking, use of XML for up marking, end-user up-marking
 - Evaluation in other domains

Literature



- Dillon, M (1991). Reader's model of text structures: the case of academic articles. *International Journal of Man-Machine Studies*, 35. 913 – 925.
- Ely, J, Osheroff, J, Ebell, M, Bergus, G, Levy, B Chambliss, M & Evans, E (1999). Analysis of wquestions asked by family doctors regarding patient care. *BMJ*, 310 (7206). 358 – 361.
- Ely, J, Osheroff, J, Gorman, P, Ebell, M, Bergus, G, Levy, B Chambliss, M, Pifer, E & Stavri, P (2000). A taxonomy of generic clinical questions: classification study. *BMJ*, 321 (7278). 429 - 432.
- Hearst, M & Plaunt, C (1993). Subtopic structuring for full length document access. *Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval*. 59 – 69.
- Jarvelin, K, Price, S, Delcambre, L M L, Nielsen, M L (2008). Discounted Cumulated Gain based Evaluation of Multiple-Query IR Sessions. *Proceedings 30th European Conference on Informaformation retrieval (ECIR), Glasgow, March 30 – April 3, 2008*
- Orlikowski, W J & Yates, J (1994). Genre repertoire: the structuring of communicative practices in organizations. *Administrative Science Quarterly*, 39. 541 – 574.
- Price, S, Delcambre, L & Nielsen, M L (2006). Using semantic components to express questions against document collections. *Proceedings International Workshop on Health Information and Knowledge Management (HIKM 2006), Arlington (VA)*.
- Price, S, Nielsen, M L, Delcambre, L & Vedsted, P (2007). Semantic components enhance retrieval of domain-specific documents. *Proceedings of the ACM Sixteenth Conference on Information and Knowledge Management (CIKM), Lisboa, November 6 - 8, 2007*.