Nineteenth International
Conference on Grey Literature

Public Awareness and Access to Grey Literature

October 23-24, 2017

National Research Council of Italy
Piazzale Aldo Moro 7, Rome

Département universitaire de Médecine générale
LIÈGE université
Unité de recherche Soins primaires et Santé

D2IM Département d'Informatique et d'Information Médicales
UNIVERSITE DE ROUEN

Indexing grey multilingual literature in General Practice in the era of Semantic Web

Marc Jamoulle[1], Melissa P. Resnick[2], Ashwin Ittoo[3], Robert Vander Stichele[4],

Elena Cardillo[5], Julien Grosjean[6], Stefan Darmoni[6], and Marc Vanmeerbeek[1]

1 Department of General Practice, University of Liège, Belgium

2 School of Biomedical Informatics, University of Texas Health Science Center at Houston, Houston, TX, USA

3 HEC Management School, University of Liège, Belgium

4 Heymans Institute of Pharmacology, University of Ghent, Belgium

5 Institute of Informatics and Telematics, National Research Council, Italy

6 Department of Information and Medical Informatics (D2IM), University of Rouen, France

Bibliography ; see text

Belgium,
Charleroi,

Home visit

Nona’s home

Photography by
Stephen Feldman, 2006

BACKGROUND

The complex world of reference of family medicine stands between anthropology and bio-sciences.

The patient doctor relationships is a heavy consumer and producer of health Information.

Productions of General Practicioners (GPs) are a silo of grey literature.

Access to information is always difficult at the point of care

There is a need for a specific indexing system, also fit for automatic coding.

BACKGROUND

Automated Question & Answer system

Example found in the publication;

John has lung cancer and has been treated with carboplatin which is known for toxicology adverse effects.

I would like to find literature and reference related to such events for the specific drug.

Sfakianaki, P., Koumakis, L., Sfakianakis, S., Iatraki, G., Zacharioudakis, G., Graf, N., ... Tsiknakis, M. (2015). Semantic biomedical resource discovery: a Natural Language Processing framework. *BMC Medical Informatics and Decision Making*, *15*(1), 77. https://doi.org/10.1186/s12911-015-0200-4

Same patient seen by his GP;

John, a Nigerian patient, has lung cancer and has been treated with carboplatin, which is known for toxicological, adverse effects.

He has been very sick and is no longer willing to follow treatment. He is depressed and expresses fear that spirits have invaded his soul. He has visited me as his family doctor to explain the situation.

I would like to find literature about patient knowledge, Nigeria cultural background, compliance, coordination of care, motivational interviewing and the role of the family doctor in managing patient denial.

A search between Family medicine, Computer sciences and Information sciences

Aim

- To improve annotation of grey literature in primary care,
- To facilitate indexing of congress abstracts and theses
- To improve the access and retrieval of these information artefacts from repositories.

Proposal

To build a taxonomy specific for indexing GPs grey literature, about clinical and contextual issues in Primary care, named 3CGP (Core Content Classification in GP/FM)

Methods

Seminal idea
Q-Codes
H.Lamberts

1987

1762 abstracts Wonca Eur., France, Swiss, Belgium
+
614 abstracts EGPRN (2010)

Paper and pencil
Excell
CAQDAS Atlas.ti

First work on contextual classification

2005

ICPC-2 + Q-Codes = 3CGP

Wonca meeting Paris

2007

Covilha meeting, Portugal

2013

Lille meeting

2014

Montevideo & Buenos Aire

2015

Paris CMF meeting

2016

Manchester & Roma meeting

2017

METHONTOLOGY steps

Knowledge acquisition

Formalization: Qualitative analysis

Integration : birth of 3CGP

HeTOP server and Webservice, Rouen

ICPC-2
(20 languages)

Q-Codes
(8 languages)
- Preferred term
- Category Label
- Entry term
- Definition
- Conceptual content
- Biblio
- Mapping MeSH
- Babelnet
- DBpedia

Implementation in HeTOP

Dissemination : website

Human indexing

Pedagogy
Bibliography
SBMFC
Telehealth
MGTFE

Computer use

Classifier
Annotator
NLP

Q-Codes versions

0.1
1.0
2.0
2.2
2.3
2.4
2.5

METHOD

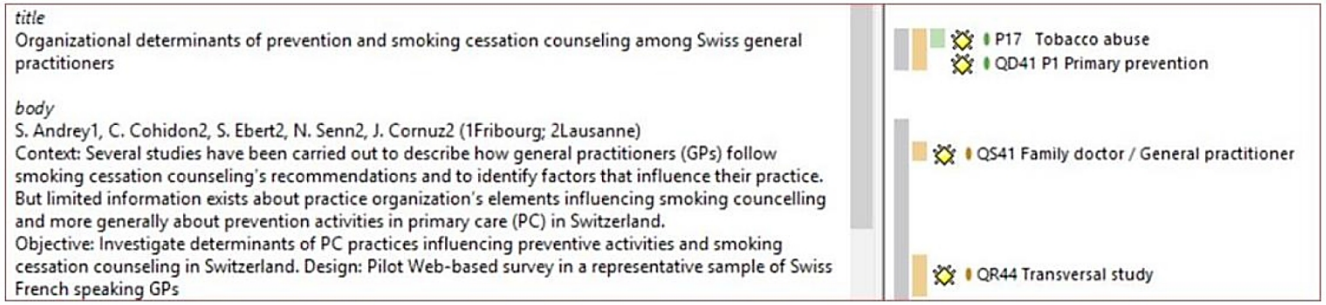
Knowledge acquisition

Qualitative Research

Paper and pencil
Excell
CAQDAS Atlas.ti

1762 abstracts Wonca Eur., France, Swiss, Belgium
+
600 abstracts EGPRN (2010)

FIGURE 1.25: Example of the coding process of an abstract (Andrey, S. et al., 2014), using the software ATLAS-ti (coding by ICPC-2 and Q-Codes, version 2.3).



Concepts Formalization

Creation of a taxonomy of 182 entries under the name of Q-Codes

Q-Codes domains and categories

- QC patient's category
  - QC1 age group
  - QC2 gender issue
  - QC3 social high risk
  - QC4 addict
  - QC5 victim of violence
  - QC6 survivor
- QD doctor's issue
  - QD1 communicator
  - QD2 doctor as carer
  - QD3 care manager
  - QD4 clinical prevention
  - QD5 complementary medicine
  - QD6 medico legal issue
  - QD7 professional image
  - QD8 work–life balance
- QE medical ethics
  - QE1 personal view
  - QE2 professional ethics
  - QE3 bioethics
  - QE4 infoethics
- QH planetary health
  - QH1 environmental health
  - QH2 biological hazard
  - QH3 nuclear hazard
- QP patient issue
  - QP1 patient safety
  - QP2 patient-centredness
  - QP3 quality of health care
  - QP4 patient perspective
  - QP5 health behaviour
  - QP6 patient participation
  - QP7 patient advocacy
- QR research
  - QR1 philosophy of science
  - QR2 epidemiology of primary care
  - QR3 research method
  - QR4 research network
  - QR5 research tools
  - QR6 expert advice
  - QR7 economics, primary health care
- QS structure of practice
  - QS1 primary care setting
  - QS2 out-of-hours
  - QS3 practice relationship
  - QS4 primary care provider
- QT knowledge management
  - QT1 teaching
  - QT2 training
  - QT3 quality assurance
  - QT4 knowledge translation
  - QT5 health data reporting

METHOD

Integration phase

ICPC + Q-Codes = 3CGP

The International Classification of Primary Care (ICPC) for indexing clinical conditions,

with adjonction of Q-Codes,

becomes the Core Content Classification of General Practice / Family Medicine

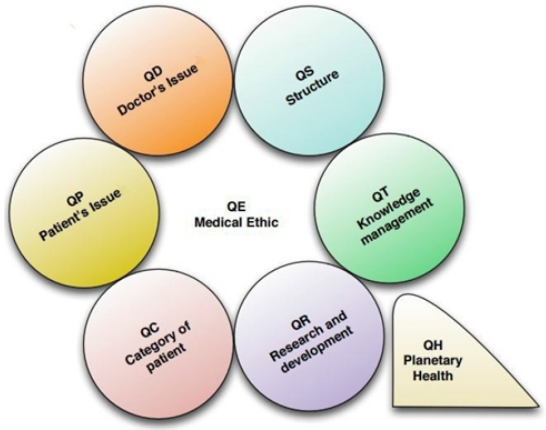
ICPC-2
(20 languages)

750 descriptors

+

Q-Codes
(8 languages)

182 descriptors



=

3CGP

http://www.ph3c.org

http://3cgp.docpatient.net/

METHOD

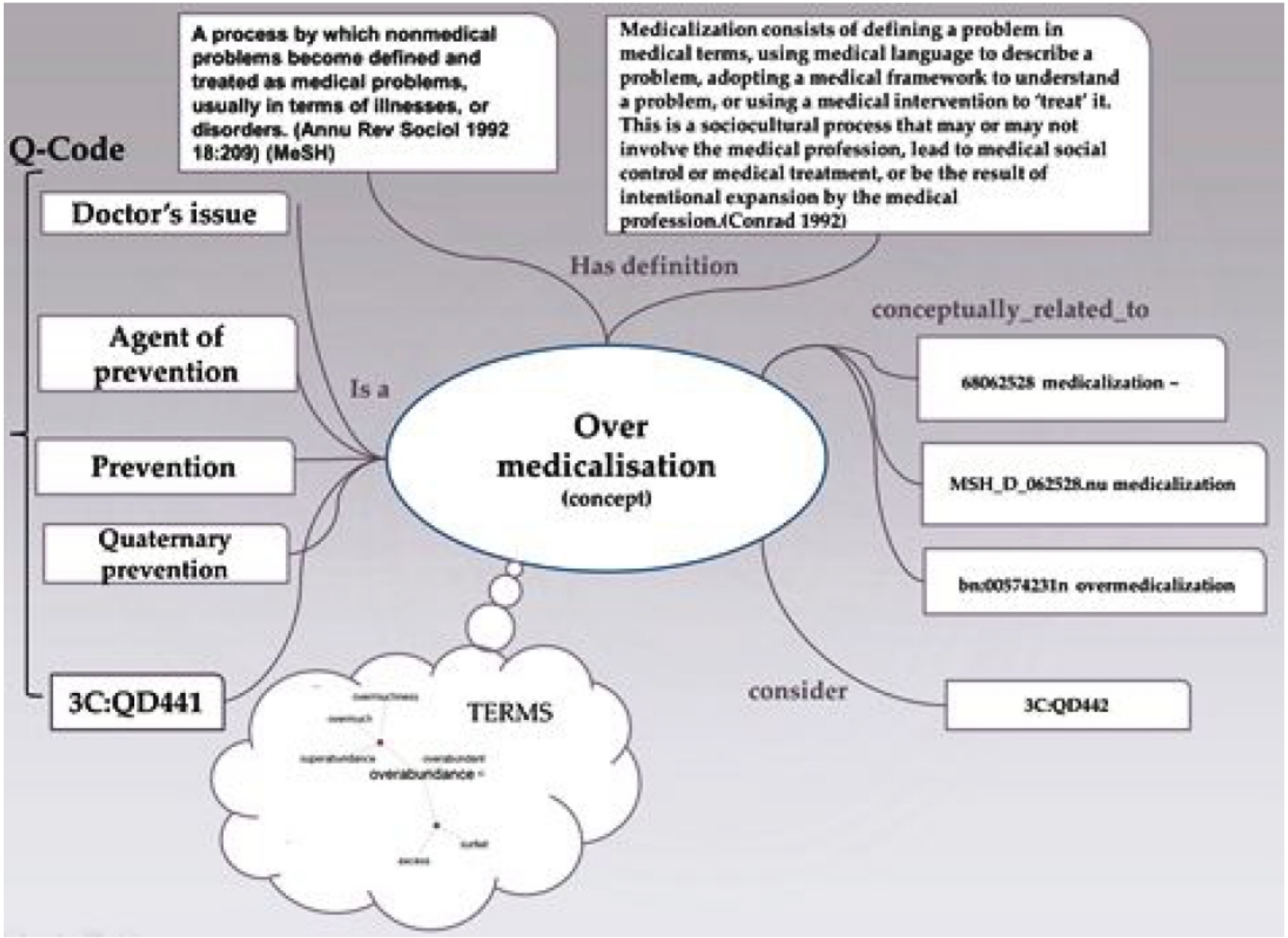
Implementation phase

HeTOP - Crosslingual Terminology Server (D2IM, Rouen , France)

Each of the 182 concepts is linked to its referents

- Q-Codes (8 languages)
  - Preferred term
  - Category Label
  - Entry term
  - Definition
  - Conceptual content
  - Biblio
  - Mapping MeSH
  - Babelnet
  - DBpedia

Data structure diagram (DSD) of a Q-Code, showing the map of concepts and their relationships (conceptual data model)



Results

We are presenting the Q-Codes, a taxonomy of contextual issues met by family doctors, complementary to ICPC, an existing Classification.

Q-Codes could be considered as a lightweight ontology ready to be used by humans but also in the semantic web context, and to be exported in Web Ontology Language (OWL)

The multilingual classes of the classification could be individually reached through Uniform Resource Identifiers (URIs).

RESULTS PRIMARY

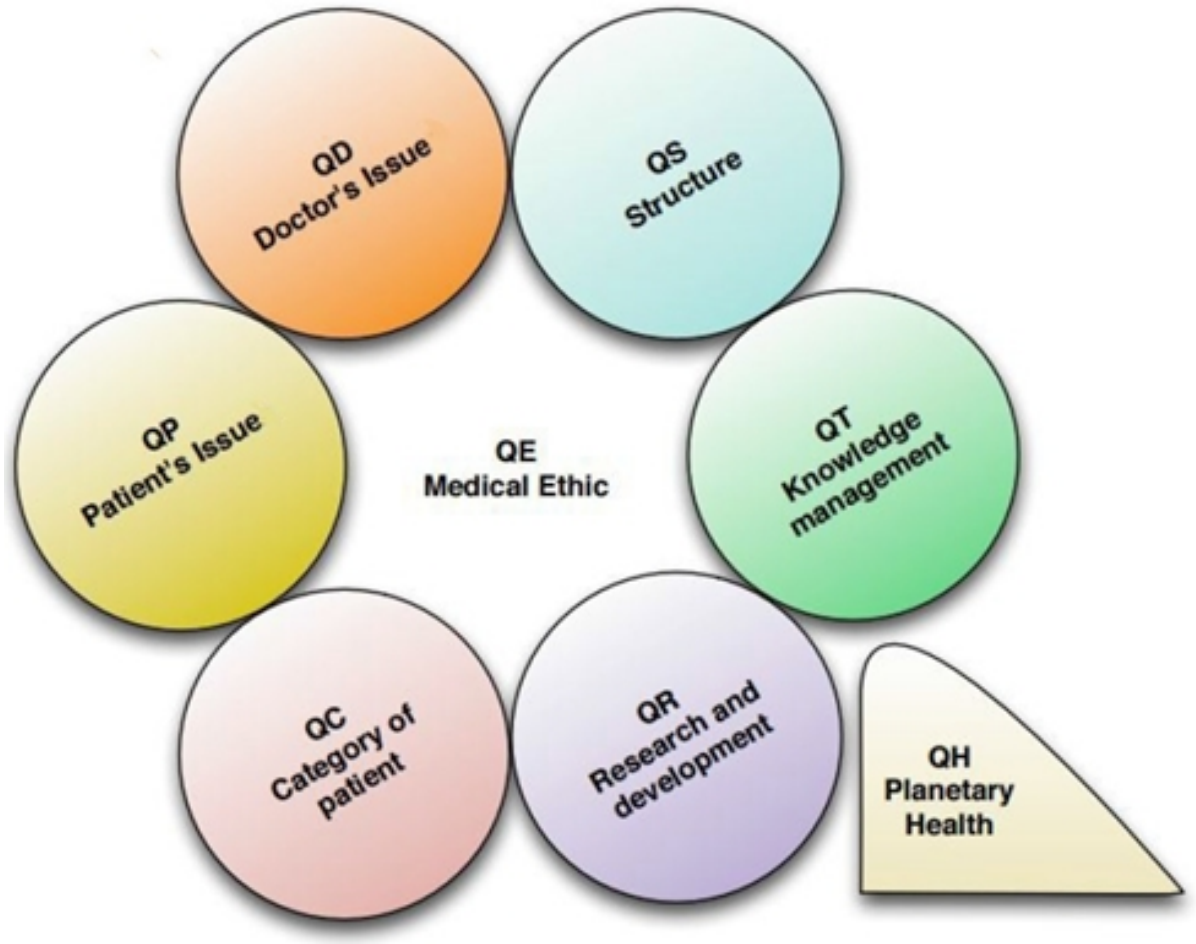
www.hetop.eu

(free access after inscription)

A taxonomy in 10 languages

The Q-Codes matrix showing the eight domains in the shape of a Q-Letter. Note that the Q's tail, which is the Planetary Health, prevents the wheel from turning endlessly.

A rag-bag QO for 'Other' has been added for suggested new codes



Terminology in book format in 6 languages

RESULTS PRIMARY

3CGP URIs

To reach the hierarchy

ICPC-2 http://www.hetop.org/hetop/?la=en&rr=CIP_C_ARBO&tab=1

ICPC-2 Process http://www.hetop.org/hetop/?la=en&rr=CIP_C_ARBOPROC&tab=1

Q-Codes http://www.hetop.eu/hetop/Q?la=en&rr=CGP_CO_Q&tab=1

To reach each rubrics

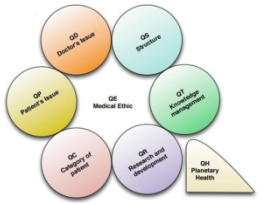
ICPC RFE and diagnosis: http://www.hetop.org/hetop/?la=en&rr=CIP_D_A01

ICPC Process http://www.hetop.org/hetop/?la=en&rr=CIP_P_30

Q-Codes http://www.hetop.eu/hetop/Q?la=en&rr=CGP_QC_QC1

- To change the language; change the ISO 639-1 for the language ;

  Ex: =en for =pt for Português (en, fr, es, pt, tr, vi, ko, nl, ka, de allowed for Q-Codes)
- To change the class; change the code at the end



RESULTS SECONDARY

Q-Codes as bibliographic tool, Belgium, 2016

l

& Portugal, 2017

Indexing of SBMFC congres and telehealth queries, Brazil, 2017

Indexing of Not@sLoc@s database, Uruguay, 2017

See the dissemination tool on http://3cgp.docpatient.net

Automated abstract classification, Belgium, 2015

Q-Codes in webProtégé

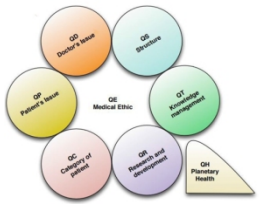
Automated French abstract annotator, France, 2017

Q-Codes in e-learning program, Vietnam, 2016

Humans: Indexing

Machines: Automated tagging

Knowledge Management Specific to GP/FM



DISCUSSION

- 3CGP, a new indexing system to be used in GP/FM available in OWL
- Allow teaching of Gp/FM
- Experiences of human indexing ongoing
- First step of automatic annotation ongoing
- Choreography by solo dancer
- Eurocentric
- Good face validity but reproductibility and Inter-doctor variation not tested

FUTURE RESEARCH

- Q Codes update needed for missed and emergent themes
- Translations (ongoing)
- Reproductibility and Interdoctorvariation
- To populate the light weight ontology
- Integration with upper level ontologies
- Automatic annotators in various language
- Standardise repositories of abstracts in GP/FM
- To take place in the distributed data era