



Care pathway records with ontologies; potential uses in medical research and healthcare

Journal:	<i>International Journal of Care Pathways</i>
Manuscript ID:	IJCP-10-021
Manuscript Type:	Conference
Date Submitted by the Author:	15-Sep-2010
Complete List of Authors:	Olive, Mark; University of the West of England, Computer Science and Creative Technologies Solomonides, Tony; University of the West of England, Computer Science and Creative Technologies
Keywords:	Care pathways, Clinical pathways, Critical pathways
Abstract:	This is one of a series of papers arising in part out of the SHARE (Supporting and structuring Healthgrid Activities and Research in Europe) project; in previous work, we have examined the use of integrated care pathways (ICPs), a fine-grained form of medical guideline including the explicit recording of any deviation, or 'variance', for research purposes. In particular, we explored how feeding the results of the analysis of variance into the development of a pathway might be an effective way of capturing 'evidence from practice'. Building on this concept, in our principal case study we propose an information system for extracting data from ICPs using ontologies and a method for inferring ICPs from other patient records, combining these with data collected for retrospective and prospective studies in preimplantation genetic diagnosis (PGD) and screening (PGS) for assisted reproduction. We also look at the problem of selecting alternatives when drug interactions occur when multiple pathways are used in parallel.

Care pathway records with ontologies; potential uses in medical research and healthcare

Mark OLIVE¹, Alison LASHWOOD² and Tony SOLOMONIDES¹

¹ Bristol Institute of Technology, University of the West of England, Bristol, UK

² Clinical Genetics, Guy's Hospital, London, UK

This is one of a series of papers arising in part out of the SHARE (Supporting and structuring Healthgrid Activities and Research in Europe) project^[1]; in previous work, we have examined the use of integrated care pathways (ICPs), a fine-grained form of medical guideline including the explicit recording of any deviation, or 'variance', for research purposes. In particular, we explored how feeding the results of the analysis of variance into the development of a pathway might be an effective way of capturing 'evidence from practice'.

Building on this concept, in our principal case study we propose an information system for extracting data from ICPs using ontologies and a method for inferring ICPs from other patient records, combining these with data collected for retrospective and prospective studies in preimplantation genetic diagnosis (PGD) and screening (PGS) for assisted reproduction. We have also looked at the problem of selecting alternatives when drug interactions might occur when multiple pathways are used in parallel.

ICPs as a source of data for research into PGD

With electronic care pathways (eICPs or ECPs) and the electronic health record (EHR) both being advanced as enhanced sources of data for research purposes, an area of interest here would be how data from both could be intelligently integrated. An example of how this might be performed emerged though our work on the EuroPGDCode project.

PGD and PGS are treatments for patients that have or are carriers of an inherited genetic disorder, which might be a chromosomal abnormality or genetic mutation, which could imply a serious health risk for any children they have. They are conducted alongside assisted reproduction, such as in-vitro fertilisation (IVF) or intracytoplasmic sperm injections (ICSI). In order to gather more information about the efficacy of PGD and PGS, members of the project based at Guy's Hospital, London produced a paediatric follow-up questionnaire, which was circulated to centres around the world. They have entered data into a system designed by project members at the University of the West of England, Bristol.

Although at an early stage, questions have already emerged. Did the parents have had any medical conditions either related or unrelated to the genetic condition in their family? Where an abnormality was picked up in a child, is this something that has been seen in either parent previously? Could it be prevalent in that family (i.e. asthma)? Where there were obstetric complications during the birth of a PGD baby, had this occurred in previous pregnancies for that couple? For example, in relation to prematurity or intra-uterine growth retardation, is this something common to the family? Did the patients have a choice of suitable embryos at the point where embryos were transferred?

While other questions could be answered by using data collected via the questionnaire, these questions require more information about the patients and previous periods of care. What would be needed is a way of extracting the relevant pieces of data from records that correspond to those previous episodes of care – this is what we will now propose.

The centralised PGD database would serve as a starting point, allowing the researcher to formulate their questions, perform exploratory queries and identify patients of interest. All records held in the database are pseudonymised. The next step is to identify which ICPs have been involved, and contain data of interest to the researcher. The patient identifiers and ICP identifiers are then sent to the distributed centres, requesting the data for those patients. For each ICP, an ontology would have been previously defined. In this context, ontologies are something that has developed from artificial intelligence research; an ontology defines all the various entities within a particular domain, what their properties and relationships are, what hierarchies exist, and so on. These ontologies are then used within decision support systems when coupled with various problem solving methods (PSMs). Hurley and Abidi^[4] have created a generic ontology for describing care pathways, incorporating temporal aspects, outcomes and variance, which serves as our starting point; Zhen et al.^[3] also demonstrated how this could be embedded into an EHR.

Where the ICP in use is recognised, data must now be extracted from the ICP record. Using an ontology to aid in this extraction allows for some variation between ICPs, rather than a formal model approach which would be less flexible. For example, Daniyal et al.^[2] demonstrated how location-specific pathways could be unified by using an ontology for prostate cancer pathways. When extracted, the data is then pseudonymised and sent back. An example of relevant

1
2
3 variance following the delivery of a PGD baby can be seen in figure 1. Where the data is held in an EHR rather than a
4 recognised ICP, an attempt must be made to match the record to an ICP. A branching care pathway or guideline for a
5 specific diagnosed condition could be expressed as a series of possible paths, each with anchor points such as referrals
6 or key test results. The EHR could be mined for these anchor points, which may correspond to periods where a
7 patient's care followed a particular guideline or pathway. The details between the anchor points can then be analysed
8 to determine if any variance was likely; at a basic level, the length of time between anchor points could be a good
9 indication that the course of care may not have been as predicted. The original pathway would then be augmented
10 with comments to indicate where complications are likely, further branching might be beneficial, etc. The formalisation
11 of the guideline or pathway would be a prerequisite for this approach, with *PROforma* being one example of a formal
12 guideline modelling language; Seyfang et al. have reported some success in easing the conversion via an intermediate
13 ontology^[5]. Where a match is made, the care pathway ontology would be employed to extract the relevant data. In
14 order to conduct a case-control study, records for controls could also be selected in a similar way at each site.

15
16
17 As mentioned above, we are also looking at the possibility of using an ontology of drug interactions to help with
18 situations where patients are following multiple ICPs, and the combination of drugs might lead to adverse reactions.
19 Examples already exist where ontologies have been employed to identify possible problems with drug interactions, but
20 these are seem limited to simple notification of the potential problem^[6]. Formal guideline modelling, using *PRODIGY* or
21 *PROforma* for example, can include components for drug selection accounting for interactions and contraindications,
22 and based on previous use^[7]. Using existing data on drug interactions, such as those collected by the British National
23 Formulary (BNF), ontology-based reasoning could also look at the severity and nature of the interaction. Where
24 variance data includes the selection of alternative drugs, and is linked to an outcome, such data could also be utilised.

25
26 Grid or cloud computing is essentially an evolution of distributed computing, where many computers are used in
27 parallel, perhaps over many sites, some for storage, some for processing, etc. The use of a grid or cloud would seem to
28 be a logical approach here, given in part that the ICP data is distributed, ICPs (and control over them) vary at a local
29 level, and the potentially computationally intensive task of extracting the required data. As the data involved is in part
30 from ICP records, the subsequent analysis could also include the identification of missing data of interest that could be
31 collected (either in the form of coded variance, or additional items that should be recorded), and any branches in the
32 pathway that might be beneficial.

33 34 **Conclusion**

35
36 The system outlined above involves integrating information in a variety of forms; retrospective and audit data collected
37 via the questionnaire and stored in a central repository, highly structured and fine grained data from care pathways,
38 stored at remote sites, and the semi-structured notes of an EHR (albeit mapped to a care pathway). The sensitive
39 nature of the data requires extensive pseudonymisation, which may not be trivial given the level of detail involved.

40
41 Our previous work involved examining ICPs as vehicles for conducting research and collecting 'evidence from practice'
42 in the form of variance records. This paper highlights another example of how ICPs might be exploited for research
43 purposes, and how data from both ICPs and EHRs might be intelligently combined through a process involving the
44 mapping of EHR data onto an ICP model, and the use of an ontology to aid in the subsequent extraction of that data.
45 The use of generic care pathway ontologies rather than specific PGD ontologies means this approach has fairly wide
46 applicability.

47 48 **References**

- 49
50 [1] Andoulsi I, Blanquer I, Breton V et al. (2008). *SHARE the Journey: A European Healthgrid Roadmap*. European Commission,
51 Directorate General Information Society and Media. ISBN: 978-92-79-09668-6.
- 52 [2] Daniyal A, Abidi S R & Abidi S S R (2009). *Computerizing Clinical Pathways: Ontology-Based Modelling and Execution*. Medical
53 Informatics in a United and Healthy Europe, IOS Press.
- 54 [3] Zhen H, Jing-song L, Hai-yan Y et al. (2009). *Modelling of Clinical Pathways Based on Ontology*. IEEE International Symposium on IT
55 in Medicine and Education (ITIME).
- 56 [4] Hurley K & Abidi S S R (2007). *Ontology Engineering to Model Clinical Pathways: Towards the Computerization and Execution of*
57 *Clinical Pathways*. IEEE International Symposium on Computer-Based Medical Systems (CBMS).
- 58 [5] Seyfang A, Miksch S, Marcos M et al. (2006). *Bridging the Gap between Informal and Formal Guideline Representations*. European
59 Conference on Artificial Intelligence (ECAI-2006).
- 60 [6] de Clercq P, Hasman A et al. (2001). *The application of ontologies and problem-solving methods for the development of shareable*
guidelines. *Artificial Intelligence in Medicine*, Vol 22, Elsevier Science B.V.
- [7] Peleg M et al. (2003). *Comparing computer-interpretable guideline models: a case study approach*. *Journal of the American Medical*
Informatics Association, Vol 10.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

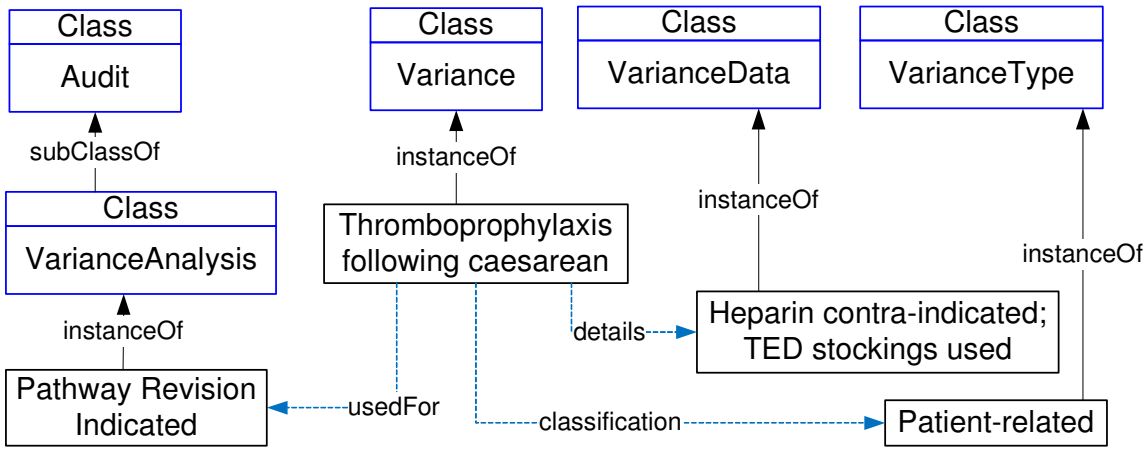


Figure 1. An example of relevant variance that might occur following the delivery of a PGD baby, using the generic care pathway ontology defined by Hurley & Abidi^[4]

Proof