Using XML and XSLT for flexible elicitation of mental-health risk knowledge

C. D. Buckingham[1,3], A. Ahmed,[1] and A. E. Adams[2]

[1] University of Aston, Aston Triangle, Birmingham B4 7ET, UK

[2] University of Warwick, Coventry CV4 7AL, UK

[3] Corresponding author: c.d.buckingham@aston.ac.uk

Keywords: XML, XSLT, knowledge engineering, mental-health risk assessment, mind maps

Runnning head: Eliciting mental-health risk knowledge

Abstract

Background: Current tools for assessing risks associated with mental-health problems require assessors to make high-level judgements based on clinical experience. This paper describes how new technologies can enhance qualitative research methods to identify lower-level cues underlying these judgements, which can be collected by people without a specialist mental-health background.
Methods and evolving results: Content analysis of interviews with 46 multidisciplinary mental-health experts exposed the cues and their interrelationships, which were represented by a mind map using software that stores maps as XML. All 46 mind maps were integrated into a single XML knowledge structure and analysed by a Lisp program to generate quantitative information about the numbers of experts associated with each part of it. The knowledge was refined by the experts, using software developed in Flash to record their collective views within the XML itself. These views specified how the XML should be transformed by XSLT, a technology for rendering XML, which resulted in a validated hierarchical knowledge structure associating patient cues with risks.
Conclusions: Changing knowledge elicitation requirements were accommodated by flexible transformations of XML data using XSLT, which also facilitated generation of multiple data-gathering tools suiting different assessment circumstances and levels of mental-health knowledge. (197 words)

(6435 words excluding abstract and references)

Tools used to assess risks associated with people who have mental-health problems require assessors to make high-level judgements that depend on considerable mental-health experience [1]. However, people exhibiting risks may present at front-line agencies (e.g. police and emergency services) where the necessary expertise for assessing them is not usually available. Rather than disappearing with their difficulties unacknowledged, a risk-screening process is required that empowers these agencies to determine the nature of the risks and whether referrals for more specialist assessments are needed. It should be based on lower-level data that are more easily recorded than high-level risk judgements.

In an ideal world, there would be a known set of factors influencing risk and associated probabilistic tools that can generate accurate predictions. Although actuarial tools exist and are often preferred because of their empirical basis [2], they fail to accommodate the transient or dynamic cues that are most important to clinical practice [3]. Clinical risk judgements depend on an individual patient's qualitative and idiosyncratic cues [4], particularly their patterns of occurrence, and risk assessments would benefit by their systematic incorporation. The problem is that people do not agree about what these cues should be [5], or how their patterns of occurrence influence assessment [6]. The motivation for this research is to find a resolution by investigating the knowledge possessed by experienced mental-health practitioners and producing a formal external representation of it.

The key question is how to break down judgements into their underlying descriptive attributes or cues. For example, most experts we interviewed mentioned depression as a high-level concept influencing risk, particularly suicide. One of its most important subcomponents is hopelessness, a lower-level, more easily assessed judgement. However, the interviews were able to reduce this even further, to issues such as having no plans, nothing to look forward to, a belief that there is no possibility for change, and so on, all of which are statements that almost anyone would be able to identify if they listened carefully enough. A risk-screening tool requires all these high-level concepts to be similarly deconstructed into lower-level patient attributes that are amenable to measurement.

This paper describes how new technologies were harnessed to traditional qualitative research methods to carry out this deconstruction of risk knowledge. In conjunction with a web-based environment, they enabled effective knowledge elicitation with multiple experts from diverse geographical locations, and helped to develop a clinical decision support system (DSS). A particular focus will be on how the eXtensible Mark-up Language, XML [7], was

used as the formal representation of expertise to facilitate collaborative use of different software and programming languages, with Extensible Stylesheet Language Transformations (XSLT) providing the flexibility for responding to inevitable changes in knowledge elicitation requirements as the research progressed.

XSLT is a language defined by the W3C to transform XML documents of a particular syntax to those of any other syntax, be they XML, html, or even plain text [8]. An XSLT stylesheet consists of a series of templates, which tell the XSLT processor exactly what to output when elements that satisfy a certain criteria are encountered in the associated XML file. All of this can be automatically achieved with most modern web browsers, which have XSLT processor libraries embedded inside them that are invoked when needed. The transformation instructions in the referenced XSLT stylesheet are applied to the XML file by the XSLT processor, and the output of this transformation is passed back to the browser to render on the screen. Hence the same base XML file can be displayed in many different ways depending on the particular requirements of the person wishing to view the information. It is a crucial property exploited by the knowledge-engineering process we used to develop our mental-health DSS.

XML is becoming increasingly important for DSSs and their development, but often with respect to enabling applications to communicate despite having different native knowledge representations [9-11]. Some systems emphasise the role of XSLT in translating XML, but with the main aim still being interoperability [12,13]. A few recognise the importance of XML and XSLT in knowledge engineering [14] and knowledge maintenance [15], especially the ease with which they can help make the process more transparent for domain experts. Our approach specifically exploits the flexibility imparted to the knowledge elicitation process itself, allowing us to alter the requirements specification dynamically without having to reprogram the elicitation tools used. As far as we are aware, ours is the only research project attempting to do this for a mental-health risk-screening DSS [16]. The purpose of the DSS will be explained first, along with the main research objectives that are the specific focus of this paper. The methods, procedures, and evolving results will then be described, followed by the conclusions and further work.

Primary objectives

The overall research programme aim is to develop the Galatean Risk Screening Tool, GRiST [16], which will be a constantly-evolving, evidence-based, world-wide web site for mental-health risk assessment. It will contain resources of three types: a database of client cues and associated risk judgements provided by practitioners as part of their clinical practice; a suite of statistical and pattern recognition tools for analysing the database; and a validated psychological model of risk assessment based on multi-disciplinary clinical expertise, which provides a full analysis of how clinicians perceive the contribution cues make to different forms of mental-health risk.

The mathematical tools generate risk predictions using processes that may not be easily understood without a strong numerate background but the psychological model is able to explain the generation of risk in terms accessible to practitioners. Together, they help bridge the gap between probabilistic information and clinical judgement, which is an important prerequisite for effective risk assessment tools [4].

The particular psychological model used to represent mental-health expertise is the galatean model of classification [17], which successfully captured clinical judgements in the related mental-health domain of psychodynamic psychotherapy [18]. It represents knowledge as a hierarchical structure and this paper explains how it was developed for risk assessment. The aim was to define a formal model of mental-health knowledge that shows how patient cues are related to each other and potential risks via a hierarchy of increasingly abstract concepts, up to the top-level risks of suicide, self harm, self neglect, harm to others, and vulnerability. Specific objectives were to:

1. elicit the knowledge structures used by individual mental-health experts when making risk assessments;

2. combine the individual structures within a single, integrated, and consensual knowledge structure, thereby identifying the low-level patient cues that need to be collected for risk assessments;

3. agree the question format and range of possible answers for each cue;

4. and generate appropriate paper-based and electronic tools for recording the patient data during or after patient assessments.

This paper will concentrate on how knowledge elicitation was enhanced by the linkage of new technologies with more traditional qualitative research methods. Analysis of the mental-health knowledge that resulted from it has been submitted for publication elsewhere.

Research design

Intelligent Knowledge-Based Systems were traditionally built using expertise from a single human source [19]. Nowadays there must be involvement of multiple experts to provide cross validation, especially in the medical field where evidence-based medicine has the ascendancy. Hence our research was designed to accommodate many experts and to obtain consensus on their collective knowledge. Individual interviews obtained the knowledge from each expert, after which web-based tasks organised along the lines of Delphi consultations [20] were conducted. These require the independent views of each participant, which are collated and sent back to the individuals, who are asked to review their input in the light of the collective sample view. Periodic focus groups were used to provide a complementary form of validation.

Methods, procedures, and evolving results

Our research was funded by an NHS New and Emerging Applications of Technology grant and Multi-Research Ethics Committee clearance was obtained. A web site for the project was established with membership facilities for all recruited participants on the expert panel. Its purpose was to maintain easy lines of communication and to conduct research remotely over the web whenever possible, a particularly important requirement for a project with panel members dissipated across the UK.

Sample

The inclusion requirement was that practitioners must be fully qualified with at least two years experience and continuing in practice for some of their time. A range of disciplines and backgrounds were recruited to provide multiple perspectives and experiences of risk assessment, encompassing academic and research areas as well as clinical practice. Most of the 46 participants who were interviewed came from psychiatric nursing (21) and psychiatry (13) but there were also some social workers, general practitioners, and psychologists. However, people were recruited on a continuous basis throughout the project and many more

were available for subsequent web tasks and focus groups, with the current panel membership consisting of over 100 clinicians and service users.

Interviews

Previous pilot work had established the general areas that needed to be covered in the interviews. These were used as prompts to stimulate further information when necessary, because the intention was to conduct open-ended interviews. The first question set the tone by asking experts to imagine they were in their normal clinical setting and assessing the risks associated with someone who was presenting with mental-health difficulties: what are the most important factors to consider?

Content analysis of interviews

Interviews were transcribed and stored as documents with numbered lines. They were then subjected to a form of content or thematic analysis [21] that aimed to identify the concepts associated with risk assessments and their constituent components. This was achieved by generating a mind map [22] for each interview using a mind-map coding template (Figure 1) that evolved as the interviews were analysed.

[PLACE FIGURE 1 ABOUT HERE]

Mind maps derived from the earlier concept maps [23] and both are useful techniques for eliciting and representing human cognition [24,25]. Mind maps are more strictly hierarchical and have a single root concept, as shown by Figure 1 with its central node labelled "risk", from which subconcepts branch out, such as history, social context, and assessment. These may be further subdivided (e.g. client episodes and family, for the history concept), so that the knowledge progresses from more general and abstract concepts in the middle to progressively more detailed notions towards the periphery. If the central risk node was picked up, the rest of the nodes would hang down like an inverted tree. This hierarchical structure is the format in which expert knowledge is to be elicited and, most conveniently, can be directly represented by XML (see below).

The mind map coding template in Figure 1 was represented using open-source software called Freemind [26] and acted as the scaffold on which to record concepts and cues in each

individual interview. Where concepts in the interview matched existing nodes of the mind map, the line number in the transcript where they were mentioned was added to the node name; if a concept was mentioned that did not match, usually because it was more individualistic and detailed than the general concepts of the template, the template was expanded to accommodate the information, again with line numbers put after the node name. When the interview had been completely coded, any template nodes without a number were removed, to leave a mind map with a common structure but individual elements. Agreement between the coding decisions of three independent researchers was better than 90% on the template categories.

[PLACE FIGURE 2 ABOUT HERE]

Once all interviews were translated into individual mind maps, each one was integrated into an emerging combined map, using a version of the template in Figure 1 that was expanded to cover more of the common knowledge revealed by the interview analysis. When a node on an individual mind map matched the combined map, or was added to the combined map if it was not already present, the expert's identification number was placed after the node name (see Figure 2) so that each node on the combined map could be linked back to the mind maps of experts who mentioned it, and from there back to the relevant lines in their interviews, thus providing a full audit trail for validation. When two researchers independently recoded the correspondence between individual mind maps and the combined one, 84% of the 125 codes associated with the transcripts in the combined map were correctly identified.

Web environment for validating the content analysis and conducting elicitation tasks

The project web site was set up to manage knowledge-elicitation tasks remotely. Dynamic exchange of information between client computers (i.e. the ones used by panel members to access the web site) and the web server (where the web site is located and managed by the research team) was achieved using PHP [27], a server-side web-scripting language. Data was stored on the server side using MySQL [28] , with graphical tasks for the panel members implemented by client-side programs, mainly using Flash [29].

The main role of the web site was enabling panel members to review their own research input and feed back on the evolving collective results, which is good practice in all qualitative research [30]. The first web task required members to comment on the mind map of their own interview and none questioned its accuracy.

Using XML to represent the mind map knowledge

Freemind stores mind maps as XML, which enables information to be transmitted in structured formats that can be customised for diverse purposes. For example, the *history* node at the top right of Figure 1 could be represented by the following XML fragment:

```
<node label="history">
<node label="client episodes">
<node label="triggers"/>
</node>
<node label="family"/>
</node>
```

where the `node` tag marks out each individual node of the hierarchy and the `label` attribute specifies the node name. A slash followed by a closing angled bracket means the node has no children (e.g. the *triggers* node), but if the slash is missing, the node is open and other nodes are nested inside (it has children, like *client episodes*). Such open nodes are closed by a matching </node> tag.

The nesting of nodes defines the hierarchical structure and the attributes of a node provide information about it. There can be any number of different attributes, which means all the information about a node that might be needed by a program processing the XML can be retrieved from the associated attributes. This property was exploited to inculcate flexibility in our knowledge engineering processes, by adding new attributes to nodes for defining how they should be displayed or transformed in ways that had not been predicted. The only proviso is that attributes have no intrinsic semantics within the XML. Neither have the tag names, and any software processing the XML must have some knowledge of the designer's preconceived meanings for them both. The XML example above is actually the format used by the programs in our project but not by Freemind, which focuses on how the XML should be displayed as a mind map. Understanding the meaning behind the Freemind XML design makes it a simple matter to translate between the different XML formats and retain the correct semantics.

Analysing the combined mind map

Depending on how many times a concept was mentioned in the interview, and which part of it, an expert's identification number (id) may appear more than once within a concept or only at the root node. The hierarchical organisation of knowledge means that all parents of a node that has an id associated with it are also deemed to be associated with the same id, because they are more general and thus encompass the specific concept. For example, referring to Figure 2, Expert 19 is associated with *most recent episode,* the top right node, and also with the ancestral nodes, *when have episodes taken place,* and *pattern of episodes,* even though the id is not explicitly recorded for those nodes.

XML is intended to be easily processed by machines and most programming languages have parsers for converting XML into data structures that can be used by the programmer. For this project, the combined mind map XML, shown in part by Figure 2, was processed using Common Lisp [31], which is eminently suited to representing hierarchical structures and effecting recursive analysis (ie, the language can extract information from tree structures easily). The Lisp program converted Freemind XML into a more convenient format and stripped the expert ids out of the node name, putting them into an attribute of the node where they were available for processing. The qualitative data has thus been transformed into the same format as the desired knowledge structure, with numerical information unambiguously defining the relationship of the structure to its interview origins. At the same time, the numbers are available to drive ensuing validation and enhancement of the knowledge, with a plethora of technologies to choose from for manipulating the XML. It represents a novel integration of qualitative research with quantitative analysis as part of knowledge elicitation.

[PLACE FIGURE 3 ABOUT HERE]

The expert ids and their association with nodes on the combined mind map was made explicit by counting all the unique ids associated with a node or its subcomponents and putting the result after the name, so that the hierarchy (or tree, as such structures are called) could be displayed with every node showing the number of experts supporting it. Figure 3 shows this for the *pattern of episodes* node in Figure 2, clearly indicating that, for example, the most recent episode and the frequency of episodes are more significant than when the first

or worst episode occurred, on the assumption that the number of different experts mentioning a node is a measure of that node's relevance. This information was used to rationalise the knowledge.

Rationalising the knowledge hierarchy

The large size of the knowledge hierarchy resulting from integrating individual experts' mind maps reflected its comprehensive representation of risk-assessment knowledge from interviews with 46 multidisciplinary mental-health experts. The complete structure (often called a tree) contained 7,210 nodes, mainly due to the high number of times many of the 1,439 uniquely-named nodes repeated across the different risks and because a generic set of concepts was maintained for experts who said they were of general importance across all risks. Out of the unique nodes, 477 were concepts (they have subcomponents or children) and 962 were leaf nodes (those without any children).

Leaf nodes such as *most recent episode* in Figure 3 define the low-level cues that will be the input data for collection when assessing a person's risks; the number of unique ones contained in the final tree will determine the number of questions to be answered by assessors. Clinicians inevitably prefer shorter assessment tools and feedback from panel experts expressed reluctance to address any that have more than between 50 and 100 questions. Our initial tree needed savage pruning but, given its size and the limited time available for experts to engage in the elicitation tasks, pruning needed to be automated, with experts only required to monitor the location of proposed cuts. The idea was to use the number of experts associated with a node as a rough measure of its importance. Concept nodes higher up the tree will tend to have more experts associated with them than those representing more detailed concepts nearer the leaf nodes (see Figure 3). Pruning by numbers will therefore naturally remove the more granular and detailed parts of the tree, as desired.

To decide the number a node requires for being spared the cut, the Lisp program was used to calculate tree sizes resulting from different thresholds and ensure that the baby was not being thrown out with the bathwater. A threshold of 5 ids for a node appeared to be keeping all the important factors and generated a more manageable structure (a similar numerically-driven approach was used by [24] when deciding which links to keep for their concept map about key success factors in a teaching hospital). The cuts in the tree now needed reviewing by the experts during a series of focus groups. It required a means of

displaying the pruning information, which the Lisp program achieved by adding the data after each node name.

Pruning the tree

[PLACE FIGURE 4 ABOUT HERE]

Software was produced for viewing the knowledge hierarchy and recording experts' views using Flash, which has a tree component for reading and displaying XML [32]. . Figure 4 shows a screen shot of the interface. The left-hand panel allows users to navigate through the tree and the right hand panel shows the immediate children of selected nodes (empty in this case because the selected node is a leaf). Various functions were written into the Flash program for viewing and annotating the tree, shown by the different buttons around the panels; Figure 4 shows a comment being added to the highlighted node.

The objective of the focus groups was to ratify the suggested cuts. The Lisp program renamed the nodes for pruning with the attention-alerting "CUT!" followed by some statistics after the name to help inform the decisions. Referring to the highlighted node in Figure 4, the first number, 3, is the number of experts mentioning this node in its particular location (it is less than the threshold of 5, which is why it has been suggested for pruning); the number in brackets is the number of experts who mentioned this node wherever it occurs in the tree (at this stage, the node is generically named to represent the first episode of any risk behaviour, not just suicide); and the numbers in square brackets state how many nodes will be removed if the cut is effected, including the number of leaf nodes (useful if the cut node is a concept that may have many subcomponents) . The important point is that the Lisp program analyses the XML to give whatever information is most useful for knowledge elicitation and passes it onto the Flash program where it can be conveniently displayed.

Conducting the focus groups

Two researchers, CDB and AEA, were present throughout all focus groups, with the Flash program projected onto a screen and the nodes displayed for reviewing the suggested cuts. CDB recorded the decisions of the focus group using the delete button to ratify the cut (a red cross obscures the node) or the comment button to record reasons why a cut should not be

made; AEA helped facilitate discussions and took additional notes. The delete and comment annotations of the tree were stored as node attributes in the XML by the Flash program and the tree was saved in the project database. After the series of focus groups reviewed the whole tree, the resulting marked up or annotated XML was checked to ensure decisions had been properly recorded for all the nodes of the tree. This is where XSLT becomes involved, because it is a language that allows one to display XML documents in precisely-specified ways.

Transforming the focus group tree using XSLT

The first use of XSLT was to display all the operations and comments put into the tree from the focus groups so that they could be easily reviewed. Any ambiguities or nodes that were missing decisions were resolved using the Flash program and the deletion commands enacted using XSLT, to transform the XML into a new, pruned, tree. The reduced hierarchy contained 3,026 nodes, with 338 unique concepts and 692 unique leaves. However, the deleted nodes were not discarded but were kept within a *help* attribute of the parent node from where they were removed, to be used in the eventual DSS for clarifying the meaning of nodes and their associated questions for assessors.

[PLACE FIGURE 5 ABOUT HERE]

Figure 5 shows the sequence of operations and transformations of data, from the interview transcripts to the output risk-screening data-gathering tools. So far, the paper has described the progression through the mind map, the tree with pruning data put into the node names, and now the pruned tree, after the pruning decisions have been enacted. The next stage is to review this pruned tree with a view to further rationalisation and reduction, by marking it up with instructions using the Flash tool and producing the fully-annotated pruned tree shown in Figure 5.

Validating and evolving the pruned knowledge hierarchy

The knowledge hierarchy was first individually validated by experts viewing their own copy of the pruned tree within the Flash tool, using all the functions as required for recording their opinions. Comments and deletes have already been explained; other buttons included

one to specify the addition of new nodes, with a description field for that node, and one to rename a node. The underlying XML data file served as the repository for the experts' input, who were able to save their annotated trees at any time, causing the XML file to be uploaded to the GRiST server.

It was a deliberate choice of the research team not to allow experts to make actual changes to the tree because of the difficulty with integrating opinions contained in trees with radically different structures. Instead, we chose to keep the one consistent tree structure and let the experts "post" their changes at the appropriate points. XSLT was then able to assimilate information about each node by displaying it as required in a browser. All the changes to each person's tree were collated in this manner and amalgamated. The collective directives were then recorded by the research team on a clean (unmarked) tree, which was validated in a second series of focus groups. The aim was to produce a new tree that will lead to the final knowledge hierarchy for the intended DSS. The flexibility of using XML and XSLT was crucial at this stage, because it became apparent that many more instructions were needed for restructuring the tree than were available via the existing Flash tool.

Enhancing the tree-changing functions

Changing requirements during knowledge engineering are inevitable because the process is not amenable to a comprehensive and watertight specification in advance [33]. Suspending the process to reprogram tools with new requirements is expensive and time consuming, even to the point of endangering successful completion. The use of XML and XSLT obviated it because the new instructions could be embedded in XML nodes using the comment box.

A set of keywords were agreed upon by the research team, and informed via the focus groups, which would indicate the operation that was later to be performed on a node. For example, it was often decided that a particular level (concept node) of the tree was redundant and that the node's children should all be attached directly to the node's parent, with the node itself disappearing. The keyword for this was *delete level,* which was written into the comment box of the node. However, these keywords were only useful if they enabled the annotated XML tree to be transmogrified automatically.

The first stage involved using XSLT to parse out the specific operations from the comments and add them to the associated nodes as proper XML attributes for subsequent processing. The resulting intermediate XML file was then processed in accordance with the attribute instructions, using the open-source *Libxslt* program [34] because it implemented

functionality specifications that were additional to the W3C XSLT 1.0 specification. It made copying a sub-tree from one location of the XML file to another location much easier than an XSLT processor limited only to the XSLT 1.0 specification.

[PLACE FIGURE 6 ABOUT HERE]

Figure 6 illustrates the relationship between the marked-up pruned tree and its transformation, as they appear in the Flash program. It shows part of the suicide risk concerned with past episodes, with the rename function enabled for the *change in pattern of episodes* node. The following XML extract shows the underlying representation of the annotated node:

```
<node renamedLabel="escalating frequency of suicide episodes"
label="change in pattern of episodes">
<node delete="delete" label="methods involved in episodes"/>
<node delete="delete" label="no reduction in seriousness"/>
<node delete="delete" label="frequency of episodes">
<node label="escalating"/> <node label="decreasing"/>
</node></node>
```

The new tree is close to the final knowledge structure but first requires questions to be added to the leaf nodes so that a data-gathering tool can be generated for use during assessments.

Defining the questions

The leaf nodes represent the information that needs to be collected by a computerised DSS, or that should inform assessments and be recorded along with any risk judgements given by the assessor, if a DSS is not being used. Either way, a form should be available for recording the data electronically or on paper, which requires each cue to be associated with a question and set of potential answer values. This was achieved by adding them as attributes to the leaf nodes, which turned out to be an additional knowledge engineering exercise, highlighting areas of the tree that were redundant or needed reorganisation. Changes were duly made and the final tree with its questions were sent out to the panel of experts for review, using the same Flash tool to record comments on the structure and PHP-driven web

forms for soliciting views on the questions. It represented the final stage in developing the knowledge hierarchy for the risk-screening DSS, producing a tree with 394 nodes in total, of which there were 124 unique concepts and 228 unique leaves.

This overall rationalisation process reduced the original knowledge structure down to a manageable hierarchy. The number of questions is still quite large, but they are not all relevant to every assessment and many depend on an affirmative response to filter questions such as "have there been any previous suicide attempts?" The screening tool thus presents a comprehensive data set for risk-screening that ought to be considered by people in front-line services who do not have a mental-health background. However, it remains too long for mental-health practitioners and it was our intention to provide an alternative assessment tool more suitable for their circumstances and experience.

Generating alternative data-gathering tools

Different assessment tools for users with varying levels of experience can be produced from the same underlying and validated knowledge structure used for the full risk-screening tool. Three levels of users were identified: level zero for those with no health background, such as the police, housing officers, or fire fighters; level one for those with a health background but not mental health, such as paramedics or accident and emergency nurses; and level two for those with a mental-health training. Level zero is the default, where all the leaf nodes are produced for the screening tool. Level one and level two nodes are defined by adding a level attribute to the node and giving it a value of 1 or 2 accordingly, accompanied by attributes that specify the question to be answered at that level and the potential responses. For example, putting a `level="2"` attribute in the depression concept node would allow a practitioner with matching expertise to assess the level of depression directly instead of recording answers about a series of cues underlying depression.

The different tools are automatically generated using XSLT. It looks for all nodes of the required level and generates the question for it when the matching level is detected; if no matching level is detected, it continues until a node with a lower level attribute is seen and produces the question associated with that one. If no level attributes are detected, it continues to the leaf nodes as normal and generates the same questions as the full risk-screening tool. A series of tools can thus be produced that suit the different backgrounds and circumstances of assessment, all linked to the same validated knowledge.

Conclusion

The goal of the research reported in this paper was to elicit the cues mental-health experts believe are important for making risk judgements and the conceptual structures that relate these cues to each other and potential risks. The process began by interviews with 46 mental-health practitioners, which were coded and integrated into a single mind map. This initial knowledge hierarchy was too large for use in a risk-screening DSS. It needed to be rationalised, refined, and reduced, which was achieved using web-based tasks, where the experts were required to annotate the XML with their opinions on the knowledge structure, using software written in Flash. The first stage was to prune the tree based on the levels of support the interview analyses showed were given to each part by the experts, ensuring the pruning did not interfere with the integrity of the knowledge. It produce a smaller tree that was subjected to rationalising and further refinement, ending with a tree that identified a set of cues that could be realistically collected during assessments. Adding questions to the cues prompted additional improvements to produce the final knowledge hierarchy. It reduced the original one from 1,439 unique nodes, of which 962 were leaf nodes (corresponding to data for collection), to a tree with just 352 unique nodes, 228 of which were leaves. It was derived through a mixture of qualitative and quantitative research methods to produce a fully-validated representation of mental-health risk-assessment knowledge. Its validity is reinforced by the triangulation of interviews, Delphi consultations, and focus groups, although not all of the experts engaged in every stage. The panel of experts numbered more than 80 by the end of this stage of research, with 46 engaging in the original interviews, between 10 and 20 conducting the different web-based tasks, and an average of 4 people participating in each of 10 focus groups.

[PLACE FIGURE 7 ABOUT HERE]

The final XML contained comprehensive information about the nodes, as shown by part of the XML for suicidal ideation in Figure 7. The main attributes define the question to ask for the associated cue, the level of expertise required for asking it, and the potential responses. Some nodes have help information for clarifying their semantics; these were automatically derived from deleted child nodes during elicitation and have not yet been translated into more

appropriate text. The `filter-q` attribute defines a question that has a binary yes/no answer, such as suicidal ideation in Figure 7 because the underlying questions are only relevant if ideation is present.

All these attributes are used by XSLT stylesheets to control how the XML is to be transformed for particular circumstances and assessors. For example, focus groups showed disagreement between how the scale values should be recorded, with some people accepting that a ten-point numerical scale was useful and others wanting an ordinal one with just three categories. Using XSLT, it is possible to give assessors the choice and generating whatever scale is preferred, without changing the underlying XML.

This fully-annotated XML knowledge hierarchy is the basis for the Galatean Risk-Screening decision support system, GRiST, currently under development [16], but it already has an important clinical role by specifying the data that should be identified during risk assessments and providing the means for collecting it. It incorporates actuarial evidence but, like the Clinical Assessment of Risk Decision Support (CARDS [35]), it is integrated with clinical knowledge. Unlike CARDS, GRiST is based on analysis of mental-health experts' knowledge structures, to produce a formal representation of how the cues combine to influence risk. Our approach is the only one to attempt this, with the goal of building a risk-screening DSS around the knowledge.

Eliciting the knowledge depended on the use of web-based technology to integrate orthodox qualitative research with the construction of intelligent knowledge-based systems. Content analysis of interview data was rendered as mind maps using XML, which then became available for processing in ways suitable both for analysing the qualitative data and for generating the hierarchical knowledge required by a DSS for risk assessment. The flexibility of XML and its rendering via XSLT, the proliferation of open-source programs, and the ease of producing bespoke software for specific manipulations, all facilitated the knowledge engineering process for obtaining the inaugural knowledge structure. It can be done remotely and means that knowledge elicitation using multiple experts is much more feasible. Nevertheless, getting the experts to carry out their tasks was a struggle [36] and focus groups turned out to be a more productive means of applying the Flash program functionality for recording views about the evolving knowledge structure.

Predicting the exact requirements of software for knowledge elicitation is unrealistic, partly due to fluidity of the engineering process, but also because experts have idiosyncratic

needs and modes of engagement. This was borne out by our research, but the crucial mitigating factor was choosing technologies such as XML and XSLT that can easily be adapted to changing circumstances. XML was the clay that bound the various programming languages and software, and XSLT was the mould that ensured the XML resource was correctly shaped for each stage of research (see Figure 5). Moulding proved invaluable when using the Flash elicitation tool, which recorded instructions for transforming the expert knowledge in ways that were not originally intended but could still be enacted using XSLT. This dynamic interaction is further exploited by the GRiST DSS itself, which generates different data-gathering tools from the same XML. Furthermore, inevitable changes to the mental-health expertise from ongoing validation can be confined to a single XML source from where all dependent tools will automatically be updated using the XSLT translation layer. This removal of redundancy to ensure integrity of GRiST's knowledge and tools reflects the design principles used in relational databases [37].

The interaction of XML and XSLT also resolved a key difficulty with qualitative research: documenting exactly how the raw data leads to results and conclusions [38]. XML both defines the evolving results of analysis and records their justification, all within the same data source. Ambiguity about how the data is transformed by the analysis is completely removed by its precise specification within an XSLT stylesheet.

Developing the GRiST DSS

The next phase of research will elicit information about how experts process cues to obtain quantified judgements of the levels of risks associated with a person's mental-health problems. This information will be incorporated within the XML knowledge structures in accordance with a psychological model of classification that formalises uncertainty processing [17]. The result will be the GRiST DSS that disseminates expert advice to professionals who do not normally have access to it, exposing many more people to early detection of their problems and the opportunity to receive appropriate assistance. GRiST will be remotely accessed through simple web-based browsers or mobile devices and its advice will help determine whether the potential risk associated with a person justifies a more detailed assessment by a specialist clinician. GRiST's potential for flexible interfaces means the specialist clinician, too, will be able to use it for conducting the assessment, thereby providing a seamless transmission of risk information that transcends disciplines and services.

References

1. Maruish ME, editor. The use of psychological testing for treatment planning and outcomes assessment. 3rd ed, vol 3. Lawrence Erlbaum, 2004.

2. Bouch J, Marshall JJ. Suicide risk: structured professional judgement. Adv Psych Treat 2005; 11: 84-91.

3. Maden A. Practical application of structured risk assessment. Brit J Psychiat 2001;178:479.

4. Holdsworth N, Dodgson G. Could a new Mental Health Act distort clinical judgement? A Bayesian justification of naturalistic reasoning about risk. J Ment Health 2003;12(5):451-462.

5. Higgins N, Watts D, Bindman, J, Slade M, Thornicroft,G. Assessing violence risk in general psychiatry. Psychiat Bull 2005;29:131-133.

6. Monahan J, Steadman HJ, Appelbaum PA, Robbins PC, Mulvey EP, Silver ER, Loren H, Grisso T. Developing a clinically useful actuarial tool for assessing violence risk. Brit J Psychiat 2000;176:312-319.

7. Harold RE, Means WS. XML in a nutshell. 2nd ed. Sebastopol,CA: O'Reilly; 2002. 613 p.

8. Kay M. XSLT 2.0 programmer's reference. 3rd ed. Wrox;2004. 960 p.

9. Dong J, Du HS, Lai KK, Wang S. XML-based decision support systems: case study for portfolio selection. Int J Inform Technol Decis Making 2004;3(4):651-662.

10. Dotsika F. From data to knowledge in e-health applications: an integrated system for medical information modelling and retrieval. Med Inform Internet 2003;28(4):231-251.

11. Taminé O, Dillman R. KaViDo−a web-based system for collaborative research and development processes. Comput Ind 2003; 42:29-45.

12. Jovanović J, Gašević D. Achieving knowledge interoperability: an XML/XSLT approach. Expert Syst Appl 2005;29:535-553.

13. Tomić B, Jovanović J, Devedžić V. JavaDON: an open-source expert system shell. Expert Syst Appl 2006;31:595-606.

14. Boegl K, Adlassnig KP, Hayashi Y, Rothenfluh TE, Leitich H. Knowledge acquisition in the fuzzy knowledge representation framework of a medical consultation system. Artif Intell Med 2004;30:1-26.

15. Del Fiol G, Rocha RA, Bradshaw RL, Hulse NC, Roemer LK. An XML model that enables the development of complex order sets by clinical experts. IEEE T Inf Technol B 2005;9(2):216-227.

16. Galatean Risk Screening Tool [Internet]. Birmingham, UK; Aston Univ - [cited 2006 Aug 8]; Available from www.galassify.org/grist/

17. Buckingham CD. Psychological cue use and implications for a clinical decision support system. Med Inform Internet 2002;27(4):237-251.

18. Buckingham CD, Birtle J. Representing the assessment process for psychodynamic psychotherapy within a computerized model of human classification. Brit J Med Psychol 1997;70: 1-16.

19. Waterman DA. A guide to expert systems. Reading, Mass: Addison-Wesley; 1986. 448 p.

20. Ayyub BM. Elicitation of expert opinions for uncertainty and risks. London: CRC Press; 2001. 302 p.

21. Neuendorf KA. The content analysis guidebook. London: Sage; 2002. 310 p.

22. Buzan T. The Mindmap book. London: BBC; 2003.

23. Novak JD. The promise of new ideas and new technology for improving teaching and learning. Cell Biol Educ 2003;2:122-132.

24. Abernethy MA, Horne M, Lillis AM, Malina MA, Selto FH. A multi-method approach to building causal performance maps from expert knowledge. Manag Acc Res 2005;16:135-155.

25. Siau K, Tan X. Improving the quality of conceptual modelling using cognitive mapping techniques. Data Knowl Eng 2005;55:343-365.

26. Freemind [Internet]. [Cited 2006 Aug 8]; Available from freemind.sourceforge.net/wiki/index.php/Main_Page

27. PHP: Hypertext Preprocessor [Internet]. [Cited 2006 Aug 8]; Available from www.php.net

28. MySQL [Internet]. [Cited 2006 Aug 8]; Available from www.mysql.com

29. Flash Professional 8 [Internet]. [Cited 2006 Aug 8]; Available from www.adobe.com/products/flash/flashpro

30. Mays N, Pope C. Qualitative research in health care: assessing quality in qualitative research. Brit Med J 2000;320:50-52.

31. CLISP – an ANSI Common Lisp [Internet]. [Cited 2006 Aug 8]; Available from www.sourceforge.net/projects/clisp

32. Jacobson D, Jacobson J. Flash and XML: A developer's guide. London: Addison-Wesley; 2002. 350 p.

33. Meyer MA, Booker JM. Eliciting and analyzing expert judgment. Philadelphia: SIAM; 2001. 459 p.

34. libxslt [Internet]. [Cited 2006 Aug 8]; Available from http://xmlsoft.org/XSLT/

35. Watts D, Bindman J, Slade M, Holloway F, Rosen A, Thornicroft G. Clinical assessment of risk decision support (CARDS): the development and evaluation of a feasible violence risk assessment for routine psychiatric practice. J Ment Health 2004;13(6):569-581.

36. Kouri P, Karjalainen-Jurvelin R, Kinnunen J. Commitment of project participants to developing health care services based on the Internet technology. Int J Med Inform 2005;74:1000-1011.

37. Date CJ. Database in depth: relational theory for practitioners. Sebastopol, CA: O'Reilly; 2005. 208 p.

38. Silverman D, editor. Qualitative research: theory, method and practice. London: Sage; 2004. 336 p.

Figure 1: Mind map template for coding interview transcripts
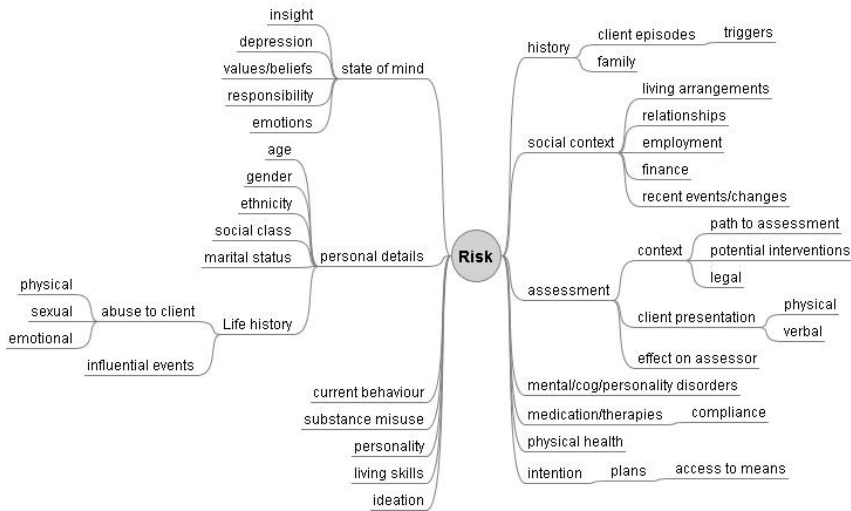
Figure 2: Pattern of suicide episodes concept showing identification numbers of experts who mentioned the particular component during their interviews
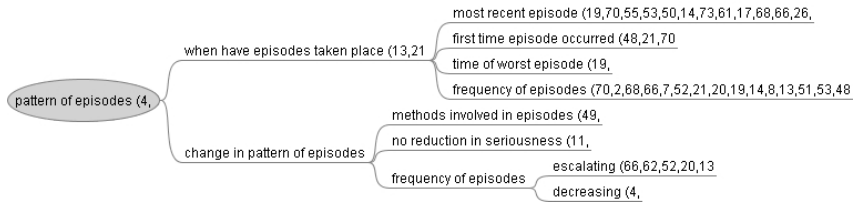
Figure 3: Part of the suicide component showing the total number of experts associated with the different elements (nodes within rounded rectangles are concepts without their internal structure displayed).
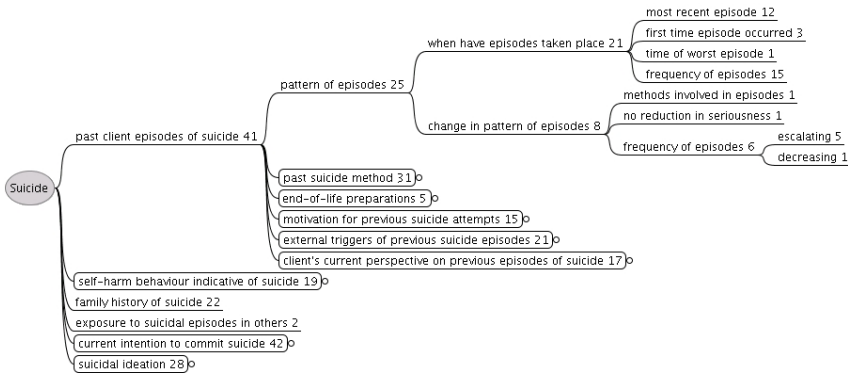
Figure 4: Flash tool showing the tree with information about where to prune it. The delete, add, rename, and comment buttons allow the tool to annotate the underlying XML document, with a comment shown being added to the highlighted node.
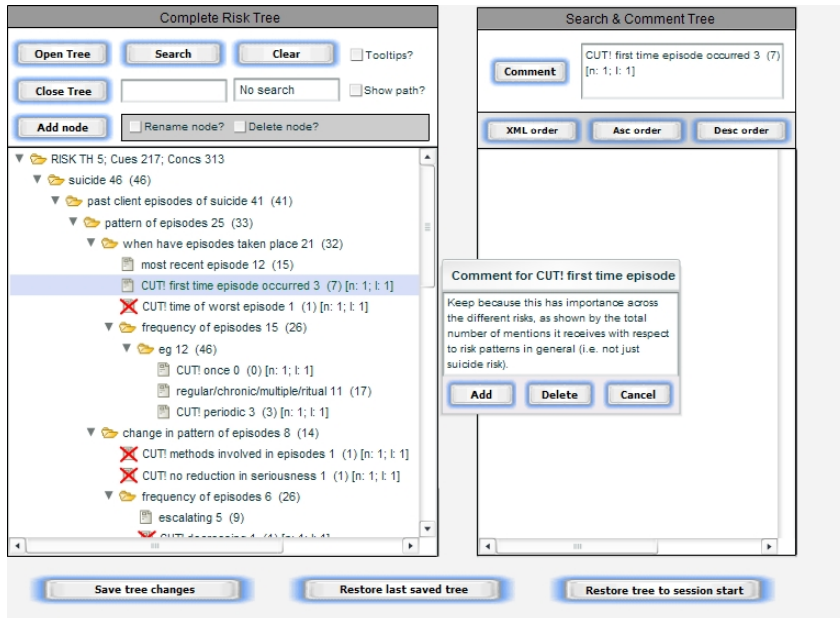
Figure 5: The sequence of XML representations and their transformations.
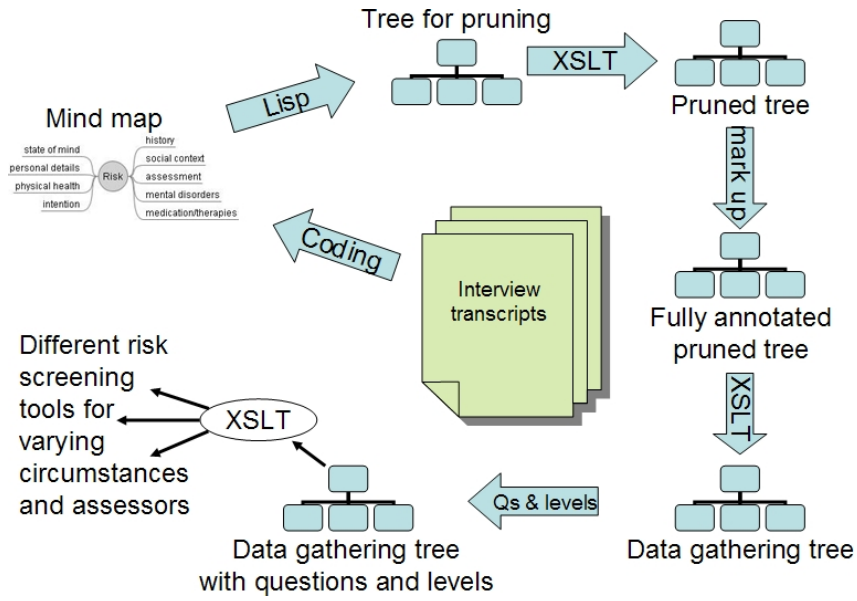
Figure 6: The fully marked up tree on the left shows how the rename (R), delete (X), and add (+) commands mark the nodes; the tree on the right shows how the attributes in the underlying XML are used by XSLT to produce a transformed tree (the initial data-gathering tree), with the highlighted node illustrating the name change.
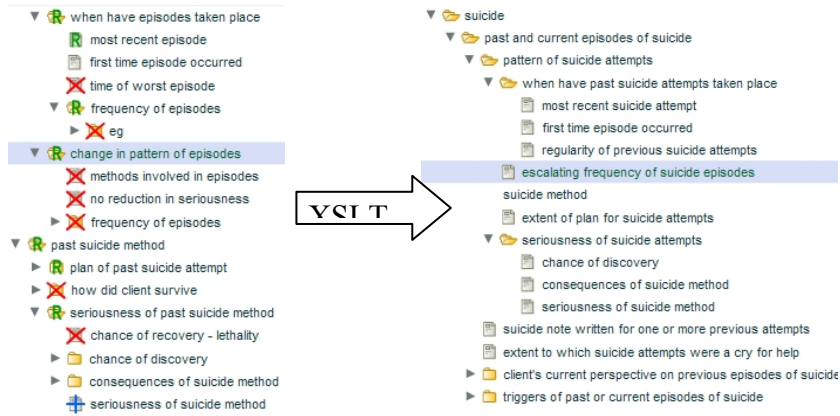
Figure 7: Part of the XML for holding information on suicidal ideation in the final knowledge hierarchy. It shows attributes holding the node names, the question to ask, the values the answer can take, the level of experience required to ask the question, and help information about the node semantics.

```
<node label="suicidal ideation"
      filter-q="Is the client having suicidal thoughts or
                fantasies?"
       question="How much do the person's suicidal thoughts/fantasies
                match those that would give you the most concern
                about suicide risk?"
      values="scale" level="1">
<node label="How much suicidal ideation is verbalised"
       question="To what extent is the person talking about suicidal
                thoughts or fantasies?" values="scale"/>
<node label="ability to control suicidal ideation"
       question="To what extent is the person able to control the
                suicidal thoughts or fantasies?" value="scale"/>
<node label="content of suicidal ideation indicates high risk"
       question="To what extent does the content of the suicidal
                thoughts or fantasies raise serious concerns about
                suicide risk?"
      values="scale" help="get away from it all ** harming others **
                       harming themselves"/>
<node label="frequency of suicidal ideation"
       question="How often do the suicidal thoughts or fantasies
                occur?"
      values="daily, less than once a day but more than once a week,
             less than once a week but more than once a month,
             less than once a month"/>
</node>
```

Acknowledgements