Innovation in Medicine and Healthcare 2014
M. Graña et al. (Eds.)
IOS Press, 2014

doi:10.3233/978-1-61499-474-9-193

Understanding Data Collection Behaviour of Mental Health Practitioners

Ali Rezaei-Yazdi [a,1] and Christopher D. Buckingham [a,2]
[a] *Department of Computer Science, Aston University, United Kingdom*

Abstract. Failure to detect patients at risk of attempting suicide can result in tragic consequences. Identifying risks earlier and more accurately helps prevent serious incidents occurring and is the objective of the GRiST clinical decision support system (CDSS). One of the problems it faces is high variability in the type and quantity of data submitted for patients, who are assessed in multiple contexts along the care pathway. Although GRiST identifies up to 138 patient cues to collect, only about half of them are relevant for any one patient and their roles may not be for risk evaluation but more for risk management.

This paper explores the data collection behaviour of clinicians using GRiST to see whether it can elucidate which variables are important for risk evaluations and when. The GRiST CDSS is based on a cognitive model of human expertise manifested by a sophisticated hierarchical knowledge structure or tree. This structure is used by the GRiST interface to provide top-down controlled access to the patient data. Our research explores relationships between the answers given to these higher-level "branch" questions to see whether they can help direct assessors to the most important data, depending on the patient profile and assessment context. The outcome is a model for dynamic data collection driven by the knowledge hierarchy. It has potential for improving other clinical decision support systems operating in domains with high dimensional data that are only partially collected and in a variety of combinations.

Keywords. clinical decision support systems, mental health, risk assessment, risk evaluation, data collection, influential variables

Introduction

The process of risk assessment and management of patients lies at the heart of health care. It is particularly challenging in the mental health sector where the symptoms, motives and deterrents are often abstract and difficult to measure. Consequently risks are often not identified in time or patients do not receive the correct level of treatment, resulting in serious adverse incidents. Early intervention usually produces better clinical outcomes [1]. A reflective analysis at two hospitals in London found that 11% of patients, had experienced some sort of adverse life events, of which 48% were judged to be

[1]Corresponding Author: Ali Rezaei-Yazdi, Aston University, Aston Triangle, Birmingham, B4 7ET; E-mail:rezaeiya@aston.ac.uk.
[2]This research was part supported by Grant SRG-0-060-11 awarded to C.D.Buckingham from the American Foundation for Suicide Prevention. The content is solely the responsibility of the authors and does not necessarily represent the official views of the American Foundation for Suicide Prevention.

preventable and 8% led to death [2]. The cost is not just to people's individual lives but also to society as a whole: £41.8 billion per year is estimated to be spent due to adverse effects of mental illness in England alone [3].

To improve detection and hence treatment of mental health problems, healthcare organisations have been increasingly turning to Clinical Decision Support Systems (CDSSs). CDSSs can improve patient health care in general [4,5] but clinicians and health organisations have often been reluctant to adopt them. Many believe they are time-consuming and require more effort compared to paper-based methods and hence are inefficient [6]. More importantly, the effectiveness of CDSSs depends on the degree to which they accommodate clinicians' cognitive workflow [7] and match their mental model of how the system should process data and operate [8]. Many fail this challenge [9], which means they do not bring key pieces of data and knowledge to the fore at the right point in the decision-making processes of the clinicians [10,7].

For these reasons, many CDSSs attempt to build their knowledge base on human expertise. The aim is to create a system that can tap into clinical expertise and use it for explaining advice in the same language. It requires identifying the facts clinicians collect and how these are used for assessing patients' health risks. It is the approach adopted for the research reported in this paper, which uses data collected by the Galatean Risk and Safety Tool, GRiST [11], developed by the second author for assessing suicide risk associated with mental health problems. The paper's focus is to exploit GRiST's model of human expertise for guiding clinicians towards the most appropriate data for an individual patient's circumstances and assessment context. The goal is a dynamic data-collection interface that responds to the emerging patient profile in synchrony with the way human experts gather evidence and reason with it.

The paper will introduce the GRiST CDSS and define the main research question with the methodology for addressing it. The experiments and results will then be discussed. The paper describes how findings can be used to create a dynamic data-collection interface that ensures the right information is collected at the right time. It ends by considering future work on how to produce a fully-working system for clinical use.

1. Brief Introduction to GRiST

GRiST is a clinical decision support system for helping mental-health practitioners assess and manage risks of suicide, self-harm, harm to others, vulnerability, and self-neglect. It is in use within many mental-health organisations in the United Kingdom and has collected half a million completed assessments so far (May, 2014). GRiST represents clinical expertise using a psychological model of classification [12] where mental-health risk knowledge is a hierarchical structure elicited in the first place from interviews with 46 experts [13]. This consensual model of risk knowledge has evolved over the years as more clinicians engaged with the research and development.

The risk nodes such as suicide are hierarchical 'trees' that are deconstructed into progressively more granular concepts (branches) such as current intention and feelings/emotions until the input data nodes or leaves of the tree are reached, such as anger.

Figure 1 shows a simplified part of the suicide risk tree. The structure provides top-down access to the patient data or cues by a series of 'filter questions' that determine whether or not a particular branch of the tree, and thus its associated leaf node cues, are

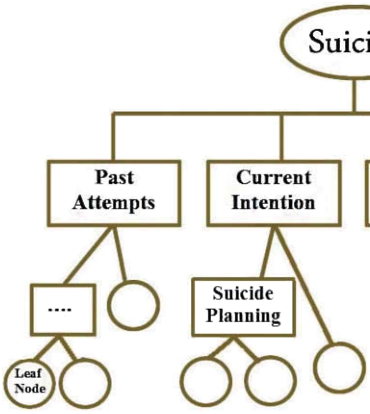

Figure 1. Part of the suicide knowledge tree.

relevant for that patient. For example, there is a high-level filter question asking whether the patient has made any past suicide attempts (as shown in Figure 1). If not, this part of the tree is immediately closed off with no further exploration. If, however clinicians say yes, then further questions open up. Clinicians carry out the assessment in this top-down fashion and at the end they are asked to provide their overall risk judgement for the patient in the form of a score ranging from 0 (no risk) to 10 (maximum risk). This paper will explore how the hierarchicy drives clinical data collection. Is there any systematic approach to how clinicians answer GRiST questions? The aim is to discover possible patterns in what data are collected, what data are missed out, and whether their mutual influence on clinical risk judgements can identify the most important data to collect next for a particular patient and assessment context.

2. Methodology and Strategy

The research investigated the answers to GRiST questions and how they relate to one another within the tree structure. The tree helps see, at higher levels, major anomalies and connections that would for example suggest special populations or patterns of data that provide insight into the how clinicians collect and use risk information. The perspective is on the semantics of the expertise and how they influences assessors' behaviour. It differs from more data-driven mathematical or statistical analyses that focus purely on the patient vectors and their relationships to risk predictions.

A software tool was developed for helping visualise and explore the relationships between variables. It identified the interesting ones and made them easy to inspect. The analysis started by looking at relationships at the top, most general, level of the tree, which involved 21 branch questions (variables) that determine which parts of the tree are relevant for further exploration. Given the high dimensionality of data (138 in suicide alone), this top-down approach addresses the combinatorial explosion. This is because it directs more detailed investigations only into the branches with interesting interactions as opposed to brute force attempts to find any relationships between all 138 leaf nodes.

In the first stage, an initial assessment of the level of dependency among variables was made using the Pearson chi-square test of independence. If variables do not show any dependency at higher level branch nodes, then it reduces the chance of any interesting relationships between data from each sub-branch lower down. If a dependency is found, the second stage explores the strength of the relationship between the variables using conditional probabilities that are presented in the form of contingency tables.

In essence, the method involved looking at how the presence or absence of one variable influences the presence or absence of another. Contingency tables show how the dependant variable, DV, is influenced by one or more independent variables, IVs. Figure 2 shows a screen shot of one such table depicting the relationship between two branch-level variables, presentation of the patient during assessment, PR, and Feelings & Emotions of the patient, FE. The cells include all combinations of the binary-valued variables, where 'yes' means the value is present and 'no' means it is absent. Each combination displays a scenario in which a relationship may be present.

Feeling & Emotion (Dependent Variable)

Presentation (Independent Variable)

P-Value < 2.2e-16	Yes	No	Total
Yes	OV: 2778 EV: 1738.21 Pos-LR: 4.36 CP: **0.82** P(B): 0.51 CP - P: **0.31**	OV: 601 EV: 1640.79 [redacted] CP: 0.18 P(�B): 0.49 CP - P: -0.31	3379
No	OV: 11834 EV: 12873.79 Neg-LR: 0.85 CP: 0.47 P(B): 0.51 CP - P: -0.04	OV: 13192 EV: 12152.21 [redacted] CP: 0.53 P(�B): 0.49 CP - P: 0.04	25026
Total	14612	13793	**GT:** 28405

Figure 2. Contingency Table

In each scenario, the tool displays various probabilities and samples, the most important of which, for this paper, are the conditional probability (CP), prior probability (P), change in probability (CP-P), and observed and expected frequency values (OV and EV respectively). Conditional probabilities show the probability of presence or absence of the DV given that the IV is present or absent. The change in prior probability of the DV shows how much influence the IV has on it. In Figure 2, presence of PR increases the probability of the presence of FE by 0.31 producing a high CP of 0.82. This suggests a relationship between tree branches that is worth exploring further.

The third stage explores the relationship between branch level variables and the clinicians' risk judgement. The method is the same as the second stage but this time investigates how presence and absence of branch level variables influence the clinical risk judgement and vice versa. Clinicians give an integer judgement between 0 (no risk) to 10 (maximum risk) and these eleven possibilities were grouped into three categories: low (score from 0-3), medium (score from 4-6), and high (score from 7 to 10).

Combining the results from stage 2 and 3, branches (i.e. variables) were identified with highest impact (causality) on other branches and on the risk judgements. A check on redundancy was also made, to determine whether and when a branch's influence is actually encompassed and explained by another branch. This is important because data collection should focus initially on information that influences risk evaluations and redundant data adds nothing to the precision. There may, of course, be other purposes for the redundant data, such as how risk should be managed, but this paper focuses on the initial evaluation. The overall results are used to produce a framework for a model that dynamically guides clinicians on what part of the knowledge tree they need to explore for rapid but effective risk evaluations.

3. Results

This section gives an example of each stage and how the results are interpreted. The outcome is an algorithm for directing data collection depending on what data has been gathered so far.

3.1. Stage 1, Test of Independence

Chi-squared tests of independence were conducted on all combinations of pairs from 21 top-level branch variables. Only pairs with a dependency having a significance level of $p < 0.01$ were investigated further. The results showed a very strong dependency for the majority of pairs but this relationship is to be expected because answering yes to one branch is likely to mean answering yes to another branch. The next stage is to determine which of the dependencies are relevant to the goal of directing data gathering.

3.2. Stage 2, Test of Causality

The software tool produced 420 contingency tables because each branch question matches each other branch question as both an IV and a DV. Thresholds of interesting relationships were set by comparing probabilities and interpreting them within the domain semantics. Two factors were important: the posterior probability of the DV and the change in prior probability of the DV. The result shows that the presence of the IV usually increases the probability of presence of the DV and, likewise, absence of IV is more associated with the absence of DV. Again, this is expected because the context of assessment influences whether or not answers are provided to the questions in general. We are interested in those relationships where the change in posterior probability is above a threshold:

1. P [DV = YES | IV = YES] > P [DV = YES] AND/OR
2. P [DV = NO | IV = NO] > P [DV = NO]

Table 1. Variables with strong relationships when both values are present. CP is the conditional probability of the dependant variable given the independent variable and CP-P is the change in probability of the DV from the prior to the conditional probability.

Independent Variable	Dependant Variable	CP	CP-P
Current Intention	Feelings & Emotion	0.86	0.34
Current Intention	Suicide Triggers	0.89	0.54
Current Intention	Suicide Ideation	0.87	0.54
Presentation	Feelings & Emotion	0.82	0.31

The variables presented in Table 1 show those that are over a conditional probability change threshold (CP-P) of 0.3. Current intention and presentation both strongly influence suicide triggers, suicide ideation, and feelings & emotion.

3.2.1. Non-additive Basis

It is possible that some DVs are not being influenced by individual IVs but by IVs in combination; the IVs are interacting to have a non-additive impact on the DV. To investigate this, the software tool compared all 1330 possible combinations[3] of two or three IVs and returned those combinations with highest influences on a DV. The result was similar to the one-to-one basis but with inevitably stronger relationships because of the multiple IVs involved. Non-additive effects were not evident but, rather, the results showed that the variables whose combination produces large CPs and changes in CPs (CP-Ps) are very similar to the variables identified in Table 1.

Thus far, the study gives evidence that there are certain branch variables within the knowledge tree which, individually or in combination, substantially increase the presence of other variables. The third stage seeks to determine the reasons for the connections and whether these should be incorporated within a dynamic data collection interface. The focus is on which data are most important for detecting risks and Stage 3 focuses on how the links between branches impact on the risk score given by clinicians.

3.3. Stage 3, Causality on Clinical Risk Judgement

Patients were divided into three risk categories as explained earlier: high, medium and low. Using the same method of contingency tables, variables were compared against each of the risk categories on a one-to-one and non-additive basis. The idea is to see whether branch relationships can guide clinicians to identify high-risk (HR) patients accurately and not falsely place them in the low-risk category. The focus was therefore on relationships linked to HR cases and Table 2 displays those with the highest value for P(HR | IV=YES).

Table 2. Influences of independent variables, IVs, on the probability of high-risk (HR) judgements. CP-P is the change in probability of HR with the IV present.

IV	**P(HR | IV)**	**CP-P**
Current Intention	0.37	0.32
Presentation	0.22	0.17
Suicide Triggers	0.14	0.09
Suicide Ideation	0.13	0.08
Presentation + Current Intention	0.51	0.48

The same variables in Table 2 are also in Table1, suggesting that those with high influence on others are also the ones that significantly influence the final risk judgement. The next task is to determine whether the variables having greatest impact on high-risk judgements do so in conjunction with others or independently.

[3]There are 1330 different ways in which 20 variables can be combined in pairs or triples. We have 21 variables in total which means for each DV we have 20 possible IVs.

3.4. Stage 4, Risk Impact Independence Test

Stage 4 explores the combinations of variables on the probability of high risk. The objective is to determine which of two singly-influential variables is the one that needs to be asked and which can be ignored because it does not give any further information. A rule was devised for testing this, as follows:

1. IF P (HR | IV & DV) $\approx$ P (HR | IV)
   AND THEN $\Rightarrow$ DV is REDUNDANT
2. P (HR | $\neg$IV & DV) $\approx$ P (HR | $\neg$IV)

where $\approx$ means the equality holds within a threshold of 0.05. The rationale is that if addition of a DV (i.e. DV = YES) does not change the probability of high risk whether the IV is YES or NO, then we can conclude that the influence on the risk can be explained entirely by the IV and the DV is redundant.

Table 3. Impact Independency Test of PR from CI

IV = Current Intention CI , DV = Presentation PR			
Condition			Threshold >0.05
P(High Risk | CI = Yes & PR = YES) = 0.51	$\neq$	P(High Risk | CI = Yes) = 0.37	**Over**
P(High Risk | CI = NO & PR = YES) = 0.04	$\approx$	P(High Risk | CI = NO) = 0.01	Within

To test the utility of the rule, it is applied to current intention and presentation because both influence the probability of high risk. Current intention is taken to be the IV because it has the strongest influence on high risk (Table 2) and what we want to know is whether presentation adds any more to what current intention is already indicating. Table 3 shows that the rule's first condition does not apply but the second one does: PR can independently increase the risk only when clinicians believe CI exists. Otherwise, when CI is set to no, PR has a negligible effect on risk. The conclusion is that presentation is not completely redundant but its influence depends on the patient's current intention.

The same rule was checked for Suicide Triggers, Suicide Ideation, and Feelings & Emotion against Current Intention and Presentation respectively. Current Intention and Presentation were taken as IVs since previous results showed that they have high causality on the other variables. The new results indicate that Suicide Triggers, Suicide Ideation and Feelings & Emotion do not influence the risk whether Current Intention is YES or NO and so are redundant as far as Current Intention is concerned: Current Intention explains the risk variation without any need for information about triggers, ideation or emotions. This is not the case for Presentation: Suicide Triggers and Suicide Ideation do have independent influence when Presentation is YES, but no independent influence when it is NO.

The results provide evidence that current intention followed by presentation are the key branches that clinicians use to make primary risk judgements. These variables are thus labelled as critical risk evaluation, CRE, variables. In contrast suicide triggers, suicide ideation and feelings & emotion are branches that, on their own, do not significantly impact the probability of high risk. However, clinicians do frequently explore and answer them, which might mean that they are important for understanding and managing the risk. Ongoing analysis is investigating this further but it is suspected that they form a second category for variables we have termed Risk Explanation variables.

3.5. Stage 5, Dealing with an Exception: Past Suicide Attempts

The final investigation stage turns the spotlight onto those exceptional cases that do not fit the rule: the samples that fall into very low probability cells of the contingency tables. For example, it was seen that P(HR | CI = NO) = 0.01, meaning when current intention is NO then the chance that clinicians assess a patient as high risk is very low, but still possible. These would be exceptional cases where patients are believed to be high risk despite having said NO to current intention. The question is why are they high risk without current intention? What makes clinicians behave differently in such cases? The hypothesis was that there could be a third variable whose presence causes these exceptional cases. Therefore, using the same visualisation tool (see Figure 2), a search was made to find the variable with highest probability of presence when Current Intention = NO and the final judgement is high risk (i.e. HR = YES).

The result showed a high presence of past suicide attempts in the exception population. In 90% of the cases when risk judgement is high and current intention is NO, there has been a history of past suicide attempts. This suggests that a history of suicide attempts should be added to the CRE variables but only if there is no current intention.

4. Final Outcome, Our Dynamic Data Collection Model

The investigations described have provided evidence for exploiting relationships between variables and risk judgements to indicate which variables are most important to ask and when they should be asked for identifying high-risk patients. The results were used to produce the skeleton of a model that helps clinicians ensure they have covered the key risk evaluation by dynamically guiding them towards the most important branches of the assessment tree. The model would start by asking clinicians to provide an answer to the current intention branch that is the most important CRE variable. If the answer is YES, then clinicians are asked to answer the second CRE variable about the patient's presentation to further assess the level of risk. However, if they say NO to current intention, then clinicians do not need to answer the questions on presentation because it is not likely to change the risk evaluation, though it may turn out to be involved in understanding and managing the risks (which is the next focus of research). Upon answering NO to current intention, clinicians are asked to provide an answer to past suicide attempts because history affects risk evaluations but only if there is no current intention. If clinicians then say NO to past suicide attempts, the system can have high confidence that the patient under assessment is not high risk: the data shows a very small probability of 0.002, without needing to explore other branches. If, however, they say YES to past suicide attempts, then although the chance of having high risk is still small, clinicians would be able to answer further questions in order to have a more accurate assessment. This might be important for those patients in the middle ground of medium risk to see if the assessor can gain more confidence that they are not high risk. Once an initial understanding of the patient health risk is established, risk-management branch questions are asked, where these are not so important for evaluating risk but, rather, how to manage it. This could be conducted at a different stage of the care pathway once the assessor is confident that the initial risk evaluation is in the right category.

5. Conclusion

For mental-health risk assessments, it is important to ask the right question a the right time. At the early risk-screening stage, data required for accurate risk evaluations are essential, before it becomes more valuable to collect data about how to manage the risks. The mental-health care pathway is long and convoluted, with each stage having its own aims and objectives. Any risk assessment tool needs to be applicable across all stages, sharing information from previous ones, but focusing on the data that is most pertinent to the task in hand. This paper has described a method of doing this by investigating the behaviour of clinicians using a semantic model of risk assessment expertise to explore probabilistic relationships.

The GRiST CDSS is able to exploit the methods described in this paper because it is explicitly based on a psychological model of mental-health risk knowledge and reasoning processes. Analysis can thus be conducted from the top-down, semantic direction as well as the bottom-up, data level. The paper has shown how the semantic approach can guide probability investigations to find relationships that would be difficult to induce with no prior knowledge. The method has already shown important results about how data can be categorised so that the right types are collected and used to guide decisions at the right points on the care pathway. The final outcome is a model for a dynamic data-gathering interface that starts by determining the primary data to be collected. Depending on those answers, if the core data indicate potential for high risk, the assessor is directed towards additional data to probe the level of risk further. Once the risk evaluation has been established, explanatory data can be requested to understand and manage the risks.

6. Future Work

Further promising work is being undertaken on detecting exceptions and how variables in different contexts, have varying influences on clinicians' decision making process. Our pilot analysis shows clinicians respond to exceptions differently and hence variables which might have been considered redundant previously, can have substantial impacts in such situations. This preliminary analysis shows that the patients' emerging profiles have a crucial influence on what data should be asked next. For example, the Stage 4 tests on redundancy show that the branch variable 'Insight and Responsibility' is redundant for the purpose of risk evaluation when statically tested against current intention (when current intention is present). However, when current intention is present but suicide ideation is absent, which is an exception, then insight and responsibility suddenly has a large influence on the probability of high risk. In other words, examining the exceptions can discover when a seemingly inert variable becomes important for risk evaluations and direct data collection accordingly. Research is continuing into detecting these situations so that we can determine where each variable should sit in the cognitive work flow of clinicians for changing patient profiles. Initial results are extremely promising for producing a fully dynamic data-collection system where the emerging profile directs the assessor to those questions that are best placed for detecting high risk.

References

[1] Swaran P Singh. Early intervention in psychosis. *The British Journal of Psychiatry*, 196(5):343–345, 2010.

[2] Charles Vincent, Graham Neale, and Maria Woloshynowych. Adverse events in british hospitals: preliminary retrospective record review. *BMJ*, 322(7285):517–519, 2001.

[3] Sainsbury centre for mental health.the economic and social costs of mental illness.policy paper 3, 2003.

[4] Amit X Garg, Neill KJ Adhikari, Heather McDonald, M Patricia Rosas-Arellano, PJ Devereaux, Joseph Beyene, Justina Sam, and R Brian Haynes. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *JAMA*, 293(10):1223–1238, 2005.

[5] Rick Goud, Mariette van Engen-Verheul, Nicolette F de Keizer, Roland Bal, Arie Hasman, Irene M Hellemans, and Niels Peek. The effect of computerized decision support on barriers to guideline implementation: a qualitative study in outpatient cardiac rehabilitation. *International Journal of Medical Informatics*, 79(6):430–437, 2010.

[6] CJ Weir, KR Lees, RS MacWalter, KW Muir, CW Wallesch, EV McLelland, and A Hendry. Cluster-randomized, controlled trial of computer-based decision support for selecting long-term anti-thrombotic therapy after acute ischaemic stroke. *QJM: Monthly Journal of the Association of Physicians*, 96(2):143–153, 2003.

[7] Dean F Sittig, Adam Wright, Jerome A Osheroff, Blackford Middleton, Jonathan M Teich, Joan S Ash, Emily Campbell, and David W Bates. Grand challenges in clinical decision support. *Journal of Biomedical Informatics*, 41(2):387–392, 2008.

[8] Ellen Kilsdonk, Linda W Peute, Rinke J Riezebos, Leontien C Kremer, and Monique WM Jaspers. From an expert-driven paper guideline to a user-centred decision support system: A usability comparison study. *Artificial Intelligence in Medicine*, 59(1):5–13, 2013.

[9] Linda WP Peute and Monique WM Jaspers. The significance of a usability evaluation of an emerging laboratory order entry system. *International Journal of Medical Informatics*, 76(2):157–168, 2007.

[10] Eta S Berner and Jacqueline Moss. Informatics challenges for the impending patient information explosion. *Journal of the American Medical Informatics Association*, 12(6):614–617, 2005.

[11] GRiST. Galatean risk and safety tool. www.egrist.org, 2014. accessed May 15th.

[12] Christopher Buckingham. Psychological cue use and implications for a clinical decision support system. *Informatics for Health and Social Care*, 27(4):237–251, 2002.

[13] Christopher D Buckingham, Ann Adams, and Chris Mace. Cues and knowledge structures used by mental-health professionals when making risk assessments. *Journal of Mental Health*, 17(3):299–314, 2008.