What Do You Want to Know?
Investigating the Information Requirements of Patient Supporters

Wendy Moncur[12], Judith Masthoff[1], Ehud Reiter[1]

*School of Computing, University of Aberdeen[1],
School of Computing, University of Dundee[2]
{wmoncur, j.masthoff, e.reiter}@abdn.ac.uk*

Abstract

There is a vast amount of data associated with any one patient. It is challenging for medical staff to understand all this data. It is even harder for a lay person, who may not even know what medical terms mean. The research project BabyTalk-Clan aims to create personalized summaries of data for a lay audience. It uses sensitive, highly-detailed clinical data relating to a patient. This includes medication given, test results, notes made by medical staff, and continuous physiological signals such as heart rate. We took a qualitative approach to knowledge acquisition for user requirements. Using interviews and a focus group within a Grounded Theory methodology, we discovered that most lay users want only a very high-level summary of the baby's state. What lay users do want is information about how the parents are coping, and what support they need. Findings were cross-validated through a questionnaire.

1. Introduction

We are exploring how to create personalized summaries of sensitive, highly-detailed clinical data for a lay audience. The summaries relate to new-born babies that are receiving specialized medical care because they are either very unwell or premature. The target audience is the parents' friends and family – what we have called the '*clan*'. There are two objectives in sending out summaries to this audience:

- to save parents time in communicating news, when their priority is the baby
- to encourage friends and family to give appropriate emotional or practical support to the parents at a difficult time.

Literature in the medical domain increasingly acknowledges the importance of keeping close supporters of a patient informed of their care, because of the vital role these people play in the patient's experience of illness [1]. As far as we are aware, this has not been addressed within the computing domain until now. In Pers4eHealth 2007, we considered parents' information dissemination choices [2]. In this paper, we explore what clan members want to know.

There is a vast amount of data available about babies in Neonatal Intensive care (NICU). Continuously monitored physiological variables (e.g. - heart rate, oxygen saturation), plus laboratory results, equipment settings and doctors' notes, create approximately 1MB of data per day [3]. Much of this data would be meaningless to anyone but a medical expert. Further, the terminology used by clinicians is highly specialized. Parents may acquire understanding of technical terms relating to their baby's condition through frequent interaction with medical staff. Clan members do not have this interaction, and are less likely to understand the implications for the baby of detailed medical information expressed in technical terms. Here, we explore what information clan members wish to receive about the baby, and how this information should be expressed.

We carried out a knowledge acquisition exercise with individuals who had previously known someone with a baby that needed extra medical care after birth. Semi-structured interviews and a focus group were used to acquire contextually-rich qualitative data. The concrete list of information components derived from this was subsequently refined via a questionnaire, which rated each item for degree of desirability in reports, and the degree of summarization required. Results indicate what personalisation is desirable from the clan members' perspective. They also show that most clan members want a very high-level summary of the baby's state, but detailed information on how the parents are coping, and what support they need.

2. Background

This work is part of the **BabyTalk** project run collaboratively between the Universities of Aberdeen and Edinburgh, and the Royal Infirmary of Edinburgh. We are creating data-to-text systems using Natural Language Generation (NLG) techniques. NLG has been used successfully in a range of systems, such as the generation of weather forecasts [4] and pun generation for children with augmentative communication needs [5]. BabyTalk seeks to produce summaries of clinical data for distinct audiences with diverse goals, from common data sources [3]:

- BT-45: 45-minute summary for medical staff.
- BT-Nurse: shift change report for nurses, covering 12 hours.
- BT-Doc: decision support for doctors, with adaptable time frame.
- BT-Family: daily summary for parents, adapted to their emotional state & gravity of baby's condition.
- **BT-Clan:** ad-hoc report for friends and family, adapted to how much information parents want to give out to individuals, and how much people want to know.

BT-Clan is unusual in considering dissemination of medical information to an audience that consists of neither medical staff nor patients. At first glance, this may seem questionable in terms of data protection and privacy. However, in the United Kingdom, patients have a right to access all of their digitally-stored medical data, and can use it as they wish. If the patient is under 16, it is the parents who have this right [6]. Parents already consider privacy. We know that parents want BT-Clan summaries to be adapted to the amount of information that they choose to give out to individual recipients [2]. Our next step has been to establish which data items are of interest to these recipients, so that we can adapt texts accordingly.

2.2 Personalisation & Social Support

Patients prefer communication to be adapted to them [7]. This can also be applied to the requirements of friends and family, in being given tailored information about the patient [2]. In our case, there are three sets of actors influencing personalisation:
- the baby, whose condition directly affects report content.
- the parents. As discussed in earlier papers, it is essential to only disseminate information according to their wishes [2, 8].
- the friends and family.

BT-Clan seeks to encourage friends and family to give **social support** to parents at a difficult time. There are three types of social support: instrumental (the provision of material aid), informational (provision of relevant information – often in the form of advice or guidance), and emotional (expression of empathy and caring, the opportunity to express feelings). Appropriate support from friends and family helps individuals to cope with stress [9]. Conversely, when social support is inappropriate, either because of the person offering it or a mismatch between the nature of the support and the stressor event, it can increase levels of stress and isolation [10]. The support that parents find appropriate cannot be inferred from the clinical data that we hold. It is unique to the individual. One mother may need instrumental support, in the form of someone collecting her older children from school whilst she is with her sick baby. Another may need the emotional support of a confidante to express her fears to. We propose that summaries be personalized to reflect parents' individual support needs.

2.3 Knowledge Acquisition & Grounded Theory

In order to create appropriate texts from all of this medical data, we need to acquire 'domain communication knowledge': how to express the data to the recipient. Lay users are unlikely to be familiar with medical terminology and may misinterpret information unless it is adapted to their level of understanding [11]. Existing approaches to acquisition of domain communication knowledge in NLG have not led to wholly successful outcomes with lay audiences: for example, personalized letters to persuade patients to stop smoking were no more successful than standard letters [12]. Approaches have included analysis of text corpora, sample text creation by domain experts, and the use of psychology models [13]. These are not appropriate for BT-Clan. We do not have a sample corpus. Nor are we aware of any existing psychology model to guide the creation of texts that will generate appropriate support. It is unsatisfactory to rely upon 'experts' – the health professionals - to deliver an understanding of lay users needs. These experts have an extensive knowledge of their own field of expertise. It can be difficult for them to see things from the lay person's perspective, and to phrase information accordingly [14].

Given the shortcomings of existing approaches to knowledge acquisition, and the lack of understanding of what information friends and family want in a health context, we chose Grounded Theory as a methodology.

Grounded Theory demands that the researcher build up a theory by gathering and analysing data, rather than testing a pre-conceived theory [15]. Although it is used mainly in social research, it has also been used in the computing domain: e.g. to explore the nature of players' enjoyment of computer gaming [16].

3. Methodology

Grounded Theory uses a technique known as 'progressive focusing', where theories, and hence research questions, evolve during the research process. We used a three-stage progressive focusing approach for knowledge acquisition, consisting of a **focus group, interviews** and a **questionnaire**. The focus group and interviews were led by a researcher experienced in group facilitation and listening skills with a perinatal population. The understanding gained in each stage informed the design of the next stage. We sought to elicit a contextually rich understanding of what information participants had received when the baby was ill, what support the participants had given to the parents at the time, and what information they would ideally have liked to receive about the parents and baby.

4. Focus Group & Interviews

4.1. Focus Group

All **participants** had experience of the child of a friend or relative receiving hospital care for serious illness. Participants were all university-educated. Gender was mixed (3 men, 1 woman), as was the age range (20-60). None of the participants had children of their own. Outcome for the babies discussed by the participants was mixed: one of the babies had died, two survived but had serious ongoing health problems and a poor life expectancy, one had made a full recovery.

During the focus group, participants were asked to outline their personal experience of receiving news about a child in hospital. Afterwards, they were given this scenario: "*Imagine that you are the close friend of someone whose baby was admitted to Neonatal Intensive Care after it was born recently*". They were asked to complete a 'card-sorting' exercise. Given a small set of cards, each showing a suggested information item, they were asked to reach a consensus on what heading to place each card under. Headings were "Essential Information", "Nice to know", "Not needed" and "Definitely don't want to know". The group was given extra cards so that they could add further items that they thought of.

4.2. Interviews

The semi-structured interviews were carried out with 4 participants (2 male, 2 female) in the age range 34-46. All participants were university-educated, and had children themselves. Outcome for the infants (3 singletons, 1 set of twins) discussed in the interviews was mixed: 4 infants had no long-term problems, 1 remains in need of further medical care and surgery.

During the interviews, participants first outlined their personal experience of receiving news about a baby in hospital. They were then asked what information they got about the parents and the baby, and what they would ideally have wanted to receive. They were asked what instrumental, informational and emotional support they gave to the parents. Finally, they were asked whether they would find automatic news updates useful, and in what way.

4.3. Analysis and Results

Transcriptions of the focus group and interviews were analysed. Categories and their properties emerged through open coding, resulting in the abstraction of themes and associated characteristics that were common across participants [16]. The results of the analysis gave us clear indications of clan members' information requirements.

Clan members tend to want sufficient information to allow planning and provision of appropriate support to the parents.

"*... one thing you might what to know is 'Is this a long term thing or a short term thing?' It would let you figure out how much of a change you need to make to your own life to properly support them*."

Focus group participant

A high-level overview of the baby's condition was preferred to a lot of detail. Clan members would provide emotional support if there was a change for the worse, or celebrate if the baby's condition improved.

"*So even with surgery ...not too interested in the details but if there is something ...that will affect the chances, I want to know that*."

Focus group participant

The exception to wanting just a summary were women who had recently had babies themselves. They wanted detailed information about how the mother was, the nature of her birth experience, the quality of medical care received (especially psychological care), and any emotional or physical repercussions of the birth for her. This was regardless of how close they were to the mother. They were aware that "excellent physical

care can …be completely negated by poor communication (with health professionals)" [17]. Their interest was "*heightened by personal experience*", giving them a "*greater ability to empathise*".

Information expressed in medical terms was not wanted. Participants felt that they would not understand it.

> "*I wouldn't have been interested.. If they'd said (*medical information*), it would have gone in one ear and out the other.*" Interviewee

Our participants were aware of the potential complexity of the parents' social network, and wished to respect it. They recognized that parents gave different people varying amounts of information. It was important to them to know who knew what information about the baby, to avoid inappropriate disclosure.

> "*My brother would be quite happy giving ME all of the information that was available. But my parents* (grandparents of the baby*), possibly not. Just a high level overview.*" Interviewee

Consistent with our previous findings [2], clan members noted that emotional distance affected information dissemination.

> "*... we were very close, and we were one of the very few to know about it… they didn't want everyone to come as well …As a family they are suffering. They don't want to talk to other people* " Focus group participant

It was also important for clan members to know when, or if, it was appropriate to visit – both to respect parents' desire for privacy, and also to meet up with other clan members, or avoid conflict with disliked ones.

In considering individual pieces of information, we found that some data items were considered "*sensitive*" – e.g – would the baby be disabled? There was no consensus on desirability of sensitive data items. 'Everyday' items, such as baby's name, date of birth and gender, were seen by some participants as non-essential. However, it was also acknowledged that these were the first things one hears when a baby is born, and that they were important for the parents to tell. A specific list of data items was derived from the outcomes of the focus group and interviews. We explored these further through a questionnaire.

5. Questionnaire

The questionnaire was designed to refine our understanding of the degree of desirability of specific data items. Of the 17 participants (7 female, 10 male), 8 had children. All were educated to university level, but employment was diverse: respondents included a speech therapist, an artist and a local government officer. There was a mixture of ages (<30 n=4; 30-45 n=9; >45 n=4). Participants were given the same scenario as the focus group: : "*Imagine that you are the close friend of someone whose baby was admitted to Neonatal Intensive Care after it was born recently*". They were then asked to specify their information requirements by choosing from options of "Essential Information", "Nice to know", "Not needed" and "Definitely don't want to know" for specific data items which had been identified in the focus group and interviews. Results are shown in Figure 1. Findings were consistent with those from the interviews and focus group. Friends and family tend to want information that enables them to support the parents effectively at a time of crisis, rather than detailed updates about the condition of the baby.

> "*I don't know the baby, but I've known the mum for years. It's her (*the mum*) that I'm worried about*". Interviewee

The results indicate what personalisation is desirable from the clan members' perspective. We have illustrated how these preferences will guide the system design in Figure 2. There is full agreement that some data items are wanted. Specifically, whether the baby's mother is in hospital herself, how the parents are coping emotionally, what practical help parents need, and a high-level overview of what is wrong with the baby. No personalisation is needed for these. However, for some data items, there was variation in whether clan members wanted them or not. For example, some respondents felt it essential to know if the baby needed surgery, others did not want this information. Personalisation will allow individual user preference on these items to be taken into account. There are two exceptions to this. Firstly, information that is customary to give in the UK when a baby is born - for example, baby's name and gender will be included in reports. These data items are not seen as sensitive. Whilst not everyone wanted to know them, it will satisfy social convention and avoid cluttering the user interface with too many personalisation choices by reporting on them. Secondly, there were some data items that some clan members wanted, but that we felt it would be inappropriate for a computer system to give out – e.g. - will the baby live or die. We will exclude these from reports.

6. Results & Future Work

Clan members have different information needs to those of medical staff and parents. Their loci of concern is the parents, and the wish to support them appropriately. The information that they seek comes from 3 sources: clinical data, local knowledge, and the parents (figure 2). Clinical data can be extracted from the shared BabyTalk resource: although only very high level summaries are required, these need to be derived from detailed domain knowledge. Local knowledge (e.g. – hospital visiting times) is static data that is straightforward to input into a system. Extracting data from the parents is a complex task. We anticipate a high degree of personalisation here, relating to what the parents' needs are, and what information they are willing to give out to whom.

In the next stage of our research, we will identify the information items that parents are happy to give out, and a user interface for parents will be prototyped. The list of data items & preferences derived in this study will be used as the starting point. Clan members want to know "How are parents coping emotionally". The the creation of an acceptable interface for parents' self-report of emotions will be an interesting challenge.

We intend to involve parents of babies currently in NICU in the next stage of research. Through the NICU, we can access a more homogenous population than we have used thus far, including those with lower education levels and more diverse social circumstances. Differences in social class, ethnicity, gender, education and occupation all influence communication and information-seeking behavior [17].

7. Discussion

As computer scientists, we can be seduced by the vast amount of data that we can manipulate, and the advances in computing which allow that data to be expressed in familiar terms. We set out to discover how individual lay users wanted high volumes of clinical data summarized. We assumed that these users would be interested in data about the baby. Yet this study found that clan members are more likely to be interested in the parents' needs than they are in the rich data stored about the sick baby. Grounded Theory coerced us to set aside preconceptions and listen to users' personal requirements, allowing fresh insights to emerge. Use of an experienced facilitator may have added to the richness of our results, as participants in the interviews and focus group talked openly about situations that were, in some cases, deeply distressing.
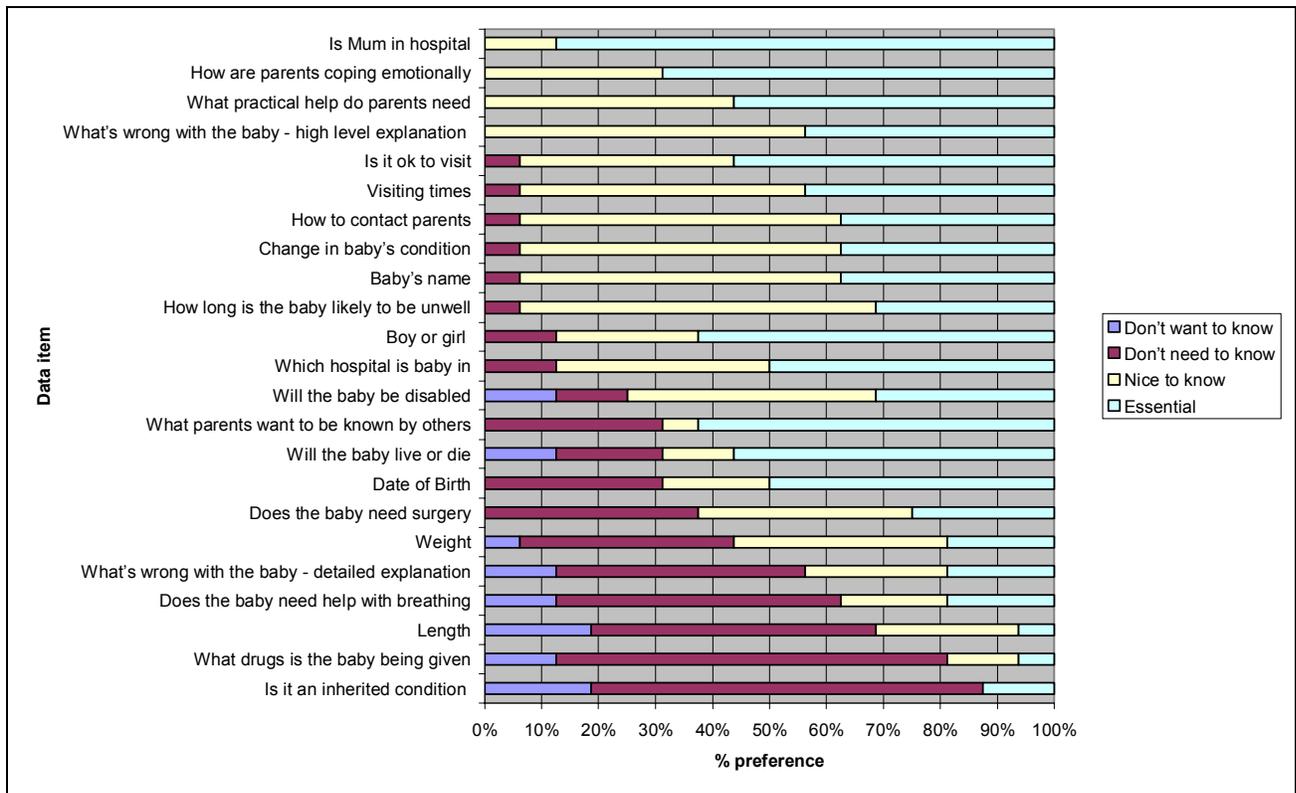

Figure 1. Information preferences of questionnaire respondents

Reported?	Data item	Source*
Socially expected	Baby's name	C
	Boy or girl	C
	Date of birth	C
	Weight	C
	Length	C
Wanted	Is Mum in hospital	C
	What's wrong with the baby (high level explanation)	C
	Change in baby's condition	C
	Which hospital is baby in	C
	Visiting times	L
	How are parents coping emotionally	P
	What practical help do parents need	P
	Is it ok to visit	P
	How to contact parents	P
Not wanted	Is it an inherited condition	C
	What drugs is the baby being given	C
Wanted by some - suitable for personal-isation	Does the baby need help with breathing	C
	What's wrong with the baby (detailed explanation)	C
	Does the baby need surgery	C
	What parents want to be known by others	P
Inappropriate to disseminate	Will the baby live or die	C
	Will the baby be disabled	X
	How long is baby likely to be unwell	X

Source: Clinical data (C), Local knowledge (L), Parents (P)

Figure 2. Effect of findings on system design

8. Acknowledgements

9. References

[1] A. O'Mara, "Who's Taking Care of the Caregiver?" *Journal of Clinical Oncology*, vol. 23, October 1. 2005, pp. 6820-6821,

[2] W. Moncur and E. Reiter. "How much to tell? disseminating affective information across a social network" in *Proceedings of the 2nd International Workshop on Personalisation for e-Health*, 2007.

[3] F. Portet, E. Reiter, J. Hunter and S. Spripada. "Automatic generation of textual summaries from neonatal intensive care data" in *Proceedings of the 11th Conference on Artificial Intelligence in Medicine*, 2007.

[4] S. Sripada , E. Reiter, J. Hunter and J. Yu., "Segmenting time series for weather forecasting" in *Applications and Innovations in Intelligent Systems* A. Macintosh, R. Ellis and F. Coenen, Eds. , 2002.

[5] G. Ritchie. "Computational mechanisms for pun generation" in *Proceedings of the 10th European Natural Language Generation Workshop*, 2005.

[6] Information Commissioner's Office, "Protecting children's personal information: ICO issues paper", 2006.

[7] A. J. Cawsey, R. B. Jones and J. Pearson, "The evaluation of a personalised health information system for patients with cancer", *User Modeling and User-Adapted Interaction 10(1),* 2000, pp. 47-72.

[8] W. Moncur. "Providing affective information to family and friends based on social networks" in *Proceedings of CHI '07: CHI '07 Extended Abstracts on Human Factors in Computing Systems*, 2007.

[9] S. Cohen, "Social relationships and health", *American Psychologist 59(8),* 2004, pp. 676-684.

[10] B. N. Uchino, J. T. Cacioppo and J. K. Kiecolt-Glaser, "The relationship between social support and physiological processes: A review with emphasis on underlying mechanisms and implications for health", *Psychological Bulletin 119(3),* 1996, pp. 488 -531.

[11] L. Alpay, P. Toussaint and B. Zwetsloot-Schonk, "Supporting healthcare communication enabled by information and communication technology: Can HCI and related cognitive aspects help?" in *Proceedings of the Conference on Dutch Directions in HCI*, 2004, pp. 12.

[12] E. Reiter, S. Sripada and R. Robertson, "Acquiring correct knowledge for natural language generation", *Journal of Artificial Intelligence Research* 18, 2003, pp. 491-516.

[13] E. Reiter, R. Robertson and L. Osman. "Knowledge acquisition for natural language generation" in *Proceedings of the First International Conference on Natural Language Generation*, 2000.

[14] R. Bromme, R. Jucks and T. Wagner, "How to refer to 'diabetes'? language in online health advice", *Appl. Cognit. Psychol*., 2005, pp. 569-586.

[15] A. Strauss and J. Corbin, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* ,2nd ed.London, United Kingdom: SAGE Publications, Inc., 1998.

[16] S. Pace, "A grounded theory of the flow experiences of Web users", *International Journal of Human-Computer Studies, vol. 60*, March 2004, pp. 327-363.

[17] J. Schott and A. Henley, *Culture, Religion and Childbearing in a Multiracial Society*, Butterworth-Heinemann, Edinburgh, Scotland, 2000.