lBirth Control, Drug Abuse, or Domestic Violence? What health risk topics are women willing to discuss with a virtual agent?

Jingjing Ren[1], Timothy Bickmore[1], Megan Hempstead[2], Brian Jack[2]

[1]College of Computer and Information Science, Northeastern University, Boston, MA, USA
{renjj, bickmore}@ccs.neu.edu
[2]Boston University School of Medicine, Boston Medical Center, Boston, MA, USA
{megan.hempstead, brian.jack}@bmc.org

Abstract. The results of a study investigating which health issues women are willing to discuss with a virtual agent are presented, focusing on 108 risks related to maternal and child health. We find that women's perceived importance of a health issue, along with general self-efficacy and comfort discussing the topic are significant predictors of whether they will address it.

Keywords: Relational agent, embodied conversational agent, preconception care, medical informatics, health informatics, social desirability.

1 Introduction

There are an increasing number of virtual agents used in the role of health counselors, for a wide range of applications, including interventions for exercise, diet, and hospital discharge. Virtual agents have also been used to screen and/or counsel individuals on potentially sensitive and stigmatizing health-related behaviors, such as drug and alcohol abuse [1] and mental illness [2]. However, it is unclear what range of health-related topics users would choose to discuss with a virtual agent, given the choice. It is also unclear what factors might drive these decisions, ranging from perceived importance of a health risk, comfort discussing a particular topic, or demographic or personality traits of the user. Such knowledge is important in designing new agent-based health interventions for particular user groups, since it would help determine whether a virtual agent should be used at all for a particular health risk and user demographic and, if a virtual agent is used, which health topics and decision factors need to be the foci of effort in intervention design.

The context of this research is an automated system we are developing to provide "preconception care" (PCC) to young African American women. These women are twice as likely to deliver a low birth weight baby and have twice the infant mortality rate compared to White women in the US. In a recent survey, 108 risks in 12 domains were identified as possible factors in determining infant health in this demographic [3] (Table 1). The domains range from substance abuse to nutrition and exercise, and individual risks span flu vaccinations to alcoholism. The majority of these risks must

be addressed well before pregnancy and traditional prenatal care, thus this area of preventive medicine is referred to as "preconception care" [4]. As just one example, folic acid should be taken at least four weeks *before* pregnancy to prevent neural tube defects.

Table 1. Domains of Health Risks Addressed in the Preconception Care System

Domain	Example Risk
Health care and programs	having health insurance
Relationships	physical or sexual abuse
Reproductive health	not using birth control
Health conditions and medicines	asthma
Genetic health history	ethnicity-based health risk
Emotional and mental health	depression
Immunizations and vaccines	need HPV vaccine
Infectious diseases	at risk for sexually-transmitted infection
Substance use	tobacco use
Nutrition and activity	overweight
Environmental issues	toxoplasmosis
Men and health care	partner does not have a physician

Users are enrolled into the PCC system by first completing a survey questionnaire that attempts to determine which of the 108 risk factors they may need to address. In our pilot work, we have discovered that the average woman in our target demographic has 23 (range 13-37) preconception risks [5]. Thus, preconception care represents an application domain in which many health behaviors need to be changed, with the set of behaviors potentially different for each user, and many of which require longitudinal counseling support. In order to help women address their preconception risks, we have developed a virtual agent (Figure 1) that counsels women over the course of a year on how to incrementally address their risks. During each session with the agent, the agent recommends that the user discuss the risk that has the highest clinical importance for preconception care, as rated by a team of family physicians. However, users are free to select any of the topics on their list of risks to discuss with the agent, and are also free to state that they feel any particular risk factor is not relevant to them.

We have recently completed a randomized clinical trial of the preconception care system, and present the results of an analysis of the topics women chose to discuss with the virtual agent. Our primary research questions are:

RQ1. What health risks do women choose to discuss with a virtual agent?

RQ2. What factors predict the risks women choose to discuss?

RQ3. What factors predict uptake on the agent's suggested risk to discuss?

2 Related Work

Many of the PCC health risk topics are very personal and may be uncomfortable for some women to talk about with others, even health professionals. Such stigma is related to social desirability bias, which is the tendency for someone to put themselves in a favorable light with respect to social norms. Kang et al. [6, 7]investigated users' socially desirable responses given different amount of self-disclosure and behavioral realism of virtual agent interviewers, finding that users disclose more with agents who have high behavioral realism and high self-disclosure. Vardoulakis explored social desirability effects in a virtual agent that queried college students about their alcohol consumption, finding that participants self-reported more frequently to a text-based interface than a virtual agent, but with interaction time accounting for the difference (the agent interface took longer to use)[8].

Discussion of personal topics, such as birth control and domestic violence, also constitutes a form of intimate self-disclosure, which generally requires some level of trust between interlocutors, according to social penetration theory [9]. Bickmore and Cassell [2] found that relational conversational strategies, such as small chat, can be used by a virtual agent, to build such a trusting relationship.

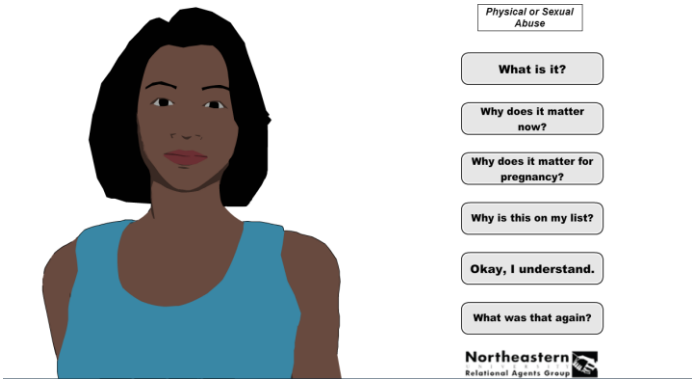

Figure 1. Virtual Agent Interface used in Preconception Care System

3 Context of the Study: The Preconception Care Virtual Agent

The Preconception Care (PCC) System is a web-based intervention that begins when a user completes a survey questionnaire to identify her PCC risks. Following this, she can conduct sessions as often as she likes to help her address her risks. The intervention is designed for a recommended weekly session over a one year period of time. Users are guided through each session by a virtual agent (Figure 1). Following the survey, the agent introduces the "My Health To-Do List" (Figure 2, Right) and lets users select the identified risks they want to discuss. The agent then describes each risk and why it is important, and offers users the opportunity to take action on it.

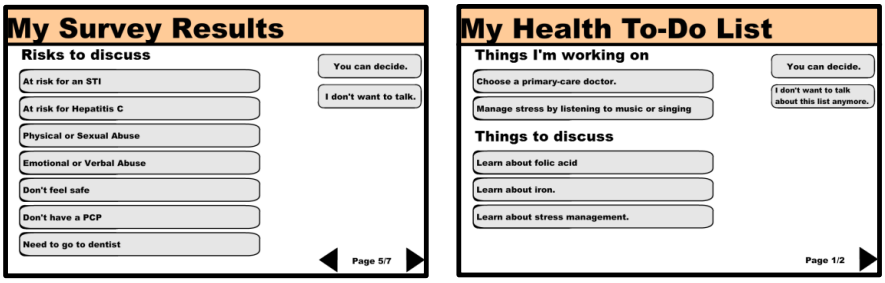

Figure 2. Sample Survey Results (Left) and "My Health To Do List" (Right)

The agent is rendered in a web-based (Adobe) Flash plugin, using the LiteBody framework [10], with speech output produced on the server with a commercial speech synthesizer. Dialogues are scripted using a custom hierarchical transition network-based scripting language. Agent nonverbal conversational behavior is generated using BEAT [11], and includes beat (baton) hand gestures and eyebrow raises for emphasis, gaze away behavior for signaling turn-taking, and posture shifts to mark topic boundaries, synchronized with speech. User input is obtained via multiple choice selection of utterances.

After a user completes her risk questionnaire her lists of risks is displayed in order of clinical importance (Figure 2, left). At this point she can either pick a risk off the list herself, or ask the agent to choose ("You can decide."), in which case the risk with the highest clinical importance is selected for discussion.

A sample introductory dialogue about a specific risk is shown in Figure 3.

```
Agent: You mentioned that you have been a victim of physical
       or sexual abuse.
User:  What is it?
Agent: Abuse is any physical or sexual harm done to you. It
       can come from any person, including family, someone you
       are dating, someone you live with, or even strangers.
User:  Why does it matter for pregnancy?
Agent: Since you are not pregnant now and are not planning on
       being pregnant soon, this is a great chance for you to
       focus on yourself and being as healthy as you can, both
       physically and emotionally. If you were abused in the
       past or are being abused now, it could make it harder to
       have a healthy baby later on. … …
User:  OK, I understand.
Agent: What best describes your attitude about getting out of a
       physically or sexually abusive situation.
User:  I plan to do this in next 6 months.
```

Figure 3. Sample Initial Dialogue for Domestic Violence Risk

4 Analysis of Health Risks Women Choose to Discuss

We recently completed a randomized controlled trial in which 100 women were recruited and randomized to either the PCC system or a non-intervention control group for six months.

4.1 Clinical Trial Methods

Participants. All 100 participants were female, self-identified as African American or Black, aged 18-34, had regular access to telephone and internet, spoke English, and were self-reported not pregnant at the time of enrollment. Participants were recruited from across the US.
Measures. Participant demographics (age, race, education), generalized self-efficacy (the confidence to overcome obstacles [12]), and health literacy (the ability to read and act on written medical information [13]), were assessed at intake. All participant actions in the PCC system were logged to a database for subsequent analysis. At the conclusion of the six-month intervention, all participants were contacted by phone and asked about each of risks identified by the Risk Assessment to determine their current status with each risk.
Protocol. Following consent, baseline assessment, and randomization, intervention participants were emailed a link to the web-based PCC survey, a Tips Sheet with instructions for using the PCC system and the study contact information. Intervention participants were emailed periodically to remind them to use the system.

4.2 Auxiliary Data

In order to determine the factors underlying decisions about whether women would discuss a particular health risk or not, we conducted an additional survey of women who were not in the clinical trial. In this survey, we asked women to assess the sensitivity and perceived importance of each of the 108 PCC health risks. We reasoned that many women may be uncomfortable discussing certain health risks, even with a virtual agent, given social desirability biases (the tendency for someone to put themselves in a favorable light with respect to social norms).
Sensitivity was assessed by asking "How comfortable would you be discussing this topic with a health professional you don't know?" on a 7-point scale (from 1="Extremely Uncomfortable" to 7="Extremely Comfortable").
Perceived Importance was assessed by asking "How important do you think this risk is for your personal health?" on a 7-point scale (from 1="Extremely unimportant" to 7="Extremely important").
The survey was distributed to 17 women (aged 22 to 29), and 15 complete data sets were obtained.

Each PCC risk was also rank-ordered in importance by a team of family physicians, based on the CDC's Select Panel on Preconception Care's publication on the clinical

content of preconception care[3]. The Panel considered not only clinical importance, but also the strength of existing evidence and efficacy of available interventions. The development of the Gabby screening questionnaire has been previously published[5]. The perceived importance of risks was significantly correlated with the clinical importance ranking, Spearman's rho=0.259, $p<0.05$. However, there were many surprising differences between clinicians' importance ranking and perceived importance by lay women. For example, use of the "withdrawal method" of birth control was ranked #5 in importance by clinicians, while lay women ranked it #93 (Table 2).

Table 2. Comparison of Clinician Importance and Perceived Importance by Lay Women of Preconception Care Risk Factors (those with most significant differences)

Risk Topic	Clinician Importance	Perceived Importance
Withdrawal Method	5	93
Multivitamin with folic acid	6	86
'Over the Counter' medicines	24	99
Household chemicals	27	101
Physical or Sexual Abuse	79	9

4.3 Results

Here we focus exclusively on the 42 women (aged 26.02+/-3.4) randomized to the PCC intervention who completed the screening questionnaire, and for whom the screening questionnaire found at least one health risk. These women completed an average of 4.19 (range 1 to 13) interactions with the virtual agent over the six months of the study. The screening questionnaire identified 23.19 (sd 6.12) risks per user and, of these, women chose to discuss an average of 6.33 (sd 7.16) risks with the virtual agent over the duration of the intervention.

RQ1. What health risks do women choose to discuss with the virtual agent?
Table 4 shows those risks women were most and least likely to discuss with the agent, once they had screened positive for the risk and agreed that it was a potential problem for them. Most women who needed HPV vaccine, were sexually active without birth control, or at risk for Hepatitis B discussed these risks with the agent, whereas none of the women who needed more Vitamin D, did not have a primary care physician, or needed more Omega-3 Fatty Acids in their diet chose to discuss these risks with the agent.

RQ2. What factors predict the risks women choose to discuss?
As we hypothesized, there are significant correlations between the perceived importance of a risk and women's likelihood of discussing it (Kendall's tau = -0.168, $p<0.05$). However, we did not find a significant correlation between the rated sensitivity (comfort discussing a risk) and women's likelihood of discussing it (Kendall's tau = -0.0148, n.s.).

Table 3. Risks Most and Least Frequently Discussed
For risks that at least 10% of women had and for likelihood of being discussed over 50% or under 5%. Columns as in **Table 5**.

Most Discussed Risks	% total	% dis /acpt	Least Discussed Risks	% total	% dis /acpt
Need HPV vaccine	22%	65%	Need more Vitamin D	20%	0%
No Birth Control	19%	64%	Don't have a PCP	14%	0%
At risk for Hepatitis B	10%	56%	Need more Omega-3 Fatty Acids	10%	0%
'Over the Counter' medicines	13%	56%	Workplace chemicals and dangers	10%	0%
Multivitamin with folic acid	26%	52%	Stress	8%	0%
Born low birth-weight or preterm	16%	50%	Personal History of Health Condition	6%	0%
Alcohol	22%	50%	Don't feel safe	6%	0%
Toxoplasmosis	19%	50%	Depression	4%	0%

Table 4. Logistic Regression to Predict Decision to Discuss Risk

Predictor	Coefficient	p
(Intercept)	-1.941	0.087
S_i	-0.165	0.088
PI_i	0.307	**0.039**
$Total_j$	-0.032	**0.015**
Age_j	-0.044	**0.046**
GSE_j	0.056	**0.001**
$REALM_j$	-0.004	0.311

In order to determine the range of factors and their relative contributions to the decision process, we performed a logistic regression on the decision to discuss each risk a woman screened positive for and agreed was potentially relevant to her. Predictors included(also see Table 5):

S_i - Sensitivity of risk i

PI_i - Perceived Importance score of risk i

$Total_j$ - Total number of risks woman j has

Age_j - Age of woman j

GSE_j - General Self-Efficacy of woman j

$REALM_j$ – Health Literacy of woman j has

Results (using the "glm" function in the R statistical analysis program) are shown in Table 4. Women's decision to discuss a risk is driven primarily by their perceived importance of the risk, but also by their general self-efficacy. The likelihood of discussing any given risk decreases with the total number of risks a woman has, indicating there may be a fixed amount of time women are willing to spend working on their health risks. The probability of discussing risks also decreases with a woman's age. The sensitivity of a risk approached significance as a predictive factor (p=.088), but

its influence was in the predicted direction: more sensitive risks were less likely to be discussed.

Table 5. Attributes of the Most Common Risks Reported in the Clinical Trial
 # total is number of women who have the risk; **# reject** is number of women who said that the screening result was incorrect; ***# discussed*** is number of women who chose to discuss the risk with the agent; ***% discussed*** = # discussed/ # total; ***%dis/acpt*** = #discussed/(#total-#reject); ***CI*** is the clinician importance rank (1=most important); ***PI*** is the perceived importance rank (from survey); ***S*** is the risk sensitivity (1-7, from survey).

Risk Topic	# total	# reject	% reject	# discussed	% discussed	% dis/acpt	CI	PI	S
Ethnicity-Based Health Risk	41	5	12.2%	12	29.3%	33.3%	51	75	3.56
Caffeine	41	10	24.4%	6	14.6%	19.4%	105	102	2.19
Listeriosis	39	15	38.5%	9	23.1%	37.5%	83	42	2.31
At risk for an STI	38	8	21.1%	7	18.4%	23.3%	75	69	4
Bad diet or food choices	35	0	0.00%	7	20.0%	20.0%	104	49	2.69
Plastic Water Bottles	33	6	18.2%	10	30.3%	37.0%	57	109	1.94
Partner needs Reproductive Life Plan	30	8	26.7%	7	23.3%	31.8%	65	105	3.06
Need more Iron	29	1	3.5%	9	31.0%	32.1%	26	85	1.59
Not been tested for an STI	27	2	7.4%	4	14.8%	16.0%	77	57	3.75
Multivitamin with folic acid	23	0	0.00%	12	52.2%	52.2%	6	86	1.59
Exercise	23	0	0.00%	2	8.7%	8.7%	68	34	2.13
At risk for Hepatitis C	23	6	26.1%	5	21.7%	29.4%	76	61	3.38
Need more Calcium	22	3	13.6%	7	31.8%	36.8%	25	81	1.59
At risk for Malaria	21	8	38.1%	6	28.6%	46.2%	71	62	2.5
Exposure to Lead	21	9	42.9%	4	19.0%	33.3%	78	83	2.13
...									

RQ3. What factors predict uptake on the agent's suggested risk to discuss?

Finally, we investigated what factors predicted when users would choose what risk they wanted to discuss vs. simply letting the agent select the risk with the highest clinical importance. The number of risks already discussed was the primary predictor of this decision. There is a general trend for women to let the agent decide more frequently the more risks they have already discussed (Figure 4). The reason for this could be that once they have already discussed the few risks that are most important to them (by choosing to discuss them), they are satisfied to let the agent choose the order of the remaining ones.

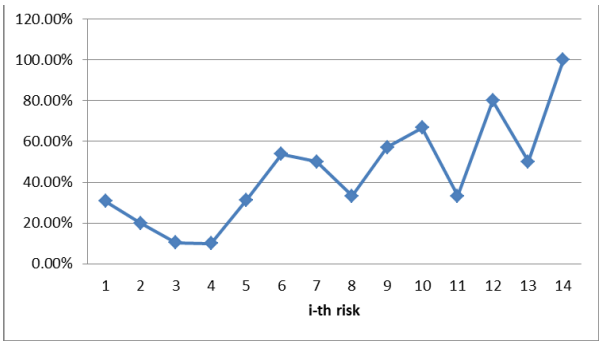

Figure 4. Percentage of Risk Decisions Abdicated to Agent, by Decision

5 Discussion, Conclusions, and Future Work

We found that women were comfortable and willing to discuss a wide range of health risks with a virtual agent, including many topics that may have significant stigma associated with them, such as domestic violence. The women in our study only discussed 6.3 of the 23.2 health risks they screened positive for with a virtual health counseling agent. However, there was a large variation in their likelihood of discussing different risks with the agent, ranging from 65% (needing HPV vaccine) to 0% (needing more Vitamin D). We found that some of the factors that contribute to this difference include the perceived importance of the risk, a woman's generalized self-efficacy, and her comfort discussing the risk.

There are several limitations to our study and analysis. The participants in the clinical trial did not complete the surveys of importance and sensitivity themselves, thus the data obtained may not match their own assessments of these factors. We also did not compare our results to decisions to discuss these risk factors with a human health counselor, which would have told us much more about women's attitudes towards virtual agents as health counselors.

We have just started another clinical trial of the PCC system in which 530 women will be randomized to the system or a non-intervention control group for a year-long period of time. Future work includes expanding the system's functionality in various ways, such as the ability for the agent to counsel women on developing a reproductive life plan. Our intent is that this effort will ultimately address the significant disparities in infant health in the US.

Acknowledgments. This work has been supported by the Agency for Healthcare Research and Quality (AHRQ), the Health Resources and Services Administration (HRSA), and the National Institute on Minority Health and Health Disparities (NIMHD). We thank the rest of the PCC team, including Daniel Schulman, Juan Fernandez, Leanne Yinusa-Nyahkoon, and Ekaterina Sadikova, Suzanne Mitchell, Paula Gardiner, Fatima Adigun, Divya Mehta and Karla Damus.

References

1. McNair, S., Checchi, K., Rubin, A., Marcello, T., Bickmore, T., Simon, S.: A Pilot Study of a Computer-Based Relational Agent to Screen for Substance-Use Problems in Primary Care. Society of Behavioral Medicine 2013 Annual Meeting, (2013)
2. Bickmore, T., Puskar, K., Schlenk, E., Pfeifer, L., Sereika, S.: Maintaining Reality: Relational Agents for Antipsychotic Medication Adherence. Interacting with Computers 22, 276-288 (2010)
3. Atrash, H., Jack, B., Johnson, K.: Where is the ‘‘w’’oman in MCH? American Journal of Obstetrics and Gynecology 199 (suppl), S259–S265 (2008)
4. Prevention, C.f.D.C.a.: Recommendations for improving preconception health and health care. Morbidity and Mortality Weekly Report 55 (RR-6), 1-23 (2006)
5. Gardiner, P., Hempstead, M.B., Ring, L., Bickmore, T., Yinusa-Nyahkoon, L., Tran, H., Paasche-Orlow, M., Damus, K., Jack, B.: Reaching Women Through Health Information Technology: The Gabby Preconception Care System Am J Health Promotion 27, 11-20 (2013)
6. Kang, S.-H., Fort Morie, J.: Users' socially desirable responding with computer interviewers. In: Conference Users' socially desirable responding with computer interviewers, pp. 229-234. ACM, (Year)
7. Kang, S.-H., Gratch, J.: Exploring Users’ Social Responses to Computer Counseling Interviewers’ Behavior.
8. Vardoulakis, L.M.: Social desirability bias and engagement in systems designed for long-term health tracking. Northeastern University Boston (2013)
9. Altman, I., Taylor, D.: *Social penetration: The development of interpersonal relationships. Holt, Rinhart & Winston, New York (1973)
10. Bickmore, T., Schulman, D., Shaw, G.: DTask & LiteBody: Open Source, Standards-based Tools for Building Web-deployed Embodied Conversational Agents. Intelligent Virtual Agents, Amsterdam, Netherlands (2009)
11. Cassell, J., Vilhjálmsson, H., Bickmore, T.: BEAT: The Behavior Expression Animation Toolkit. In: Conference BEAT: The Behavior Expression Animation Toolkit, pp. 477-486. (Year)
12. Schwarzer, R., Jerusalem, M.: Generalized Self-Efficacy scale. In: Weinman, J., Wright, S., Johnston, M. (eds.) Measures in health psychology: A user’s portfolio, pp. 35-37 (1995)
13. Davis, T., Long, S., Jackson, R., al, e.: Rapid estimate of adult literacy in medicine: a shortened screening instrument. Fam Med 25, 391-395 (1993)