Incorporating Personalization Features in a Hospital-Stay Summary Generation System

Sabita Acharya, Andrew D. Boyd, Richard Cameron, Karen Dunn Lopez, Pamela Martyn-Nemeth,
Carolyn Dickens, Amer Ardati, Barbara Di Eugenio
University of Illinois at Chicago, Chicago, IL
{sachar4, boyda, rcameron, kdunnl2, pmartyn, cdickens, aardati, bdieugen}@uic.edu

Abstract

Most of the currently available health resources contain vast amount of information that are created by keeping the "general" population in mind, which in reality, might not be useful for anyone. One approach to providing comprehensible health information to patients is to generate summaries that are personalized to each individual. This paper details the design of a personalized hospital-stay summary generation system that tailors its content to the patient's understanding of medical terminologies and their level of engagement in improving their own health. Our summaries were found to cover around 80% of the health concepts that were considered as important by a doctor or a nurse. An online survey conducted on 150 participants verified that our algorithm's interpretation of the personalization parameters is representative of that of a larger population.

1. Introduction

According to the Centers for Disease Control and Prevention, 50% of all the deaths that occur due to heart issues in the US can be prevented by making some changes in one's lifestyle [1]. Discharge notes and patient education materials that are given to patients when they are discharged from the hospital provide more information regarding the health issue and include general suggestions on how patients can take care of themselves. However, researchers from various medical fields have demonstrated that a substantial proportion of the US population find it difficult to comprehend patient education materials [2]. Patients who clearly understand their after-hospital care instructions are 30% less likely to be readmitted compared to those patients who lack this information [3][1]. Moreover, most of the health documents are developed by using "one size fits all" approach and fail to address various factors (like the patient's concerns, preferences, and ability to understand medical terms) that can play a significant role in determining the kind of information that a patient can comprehend. The Office of Disease Prevention and Health Promotion mentions that even an individual's ability to understand medical information might not be the same before and after they are confronted with a stressful situation [5]. Hence, all of these factors need to be taken into consideration while developing effective health documents for patients.

Our solution to this problem is to provide a comprehensible and personalized summary of what happened to the patient in the hospital. Since doctors and nurses focus on different aspects of patient care [6] and patients with chronic illness (like heart issues in our case) will need to continue much of the care that was provided to them in the hospital by the nurses [7], we summarize the information from both the physician and nursing documents. To the best of our knowledge, there are no existing systems that generate personalized content based on patient-specific features like their

---

[1]Lower readmission is equally important to hospitals since the Hospital Readmission Reduction Program financially penalizes hospitals that have high 30-day readmission rates [4]
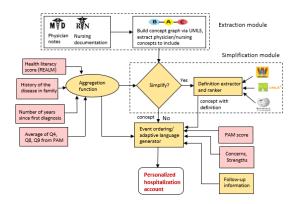
HICSS

Figure 1. The inputs, output, and workflow of our personalized summary generation system

preferences, concerns, and motivation in taking care of their health. Moreover, the combination of the four factors that guide our personalization process have never been explored before.

2. Related Work

Research on producing personalized content has been pursued for quite some time now. Some systems are able to produce different versions of the same content, e.g., generating descriptions of devices for people of different age groups [8], while others produce a single version that is aimed at some specific audience, e.g., summarizing news stories for children [9]. Mairesse and Walker [10] developed a parameterizable language generator that takes the user's linguistic style into account and generates restaurant recommendations. CrAg2 system [11] produces various linguistic styles that can lead to different perceptions of character personality. Some of the existing systems in the biomedical domain produce personalized content for patients. The seminal work by Jimison et al.[12] allows user to browse through generic information on their health issue and possible treatments. DiMarco et al.[13] assumes that a master document with annotations on when to include what content is already present and uses this information for personalizing content. The BabyTalk system [14] generates textual summaries of neonatal data for

people occupying different roles (e.g. nurses, parents) in the Neonatal Intensive Care Unit. They use handcrafted ontologies, which are very time and resource intensive to construct. Erriquez and Grasso [15] also make use of an ontology structure for generating personalized advisory messages.

3. System Workflow

This study builds upon our summary generation system and introduces personalization features into the summaries. The complete workflow of our system is shown in Figure 1. This system takes the physician and nursing documentation as its input. The nurses use structured nursing terminologies in a plan of care (POC) software called HANDS [16] for recording patient information during each shift. HANDS uses three different taxonomies to characterize the overarching elements of the nursing process: NANDA-I for nursing diagnosis [17], NOC for outcomes [18], and NIC for interventions [19]. On the other hand, the doctor's discharge note consists of the medical aspects of the hospitalization and is in free text format. When this system is deployed in real time at a certain point in the future, the physician and nursing notes would be taken from the Electronic Medical Record, while the inputs (shown in pink in Figure 1) all come from the patients and would be collected appropriately, e.g., via a tablet. In this paper, we will summarize the functioning of the Extraction module (see Section 3.1) and the Simplification module (see Section 3.2), both of which are part of our previous works [20, 21]. Details on our personalization process, including the specific inputs that the algorithm will require of the patient will be explained in Section 4.

3.1. Extraction module

The extraction module in Figure 1 represents the core of our Natural Language Generation pipeline and includes a framework for comparing the relationship between the medical terms present in the nursing documentation and the doctor's discharge note. We use an information extraction tool called MedLEE [22] for extracting all the

Table 1. Features that are extracted for modeling complexity

Category	Features
Lexical features	Number of vowels, consonants, prefixes, suffixes,letters, syllables per word
Counts of each POS type	Number of nouns, verbs, adjectives, prepositions, conjunctions, determiners, adverbs, numerals (using Stanford parser)
Vocabulary based	Normalized frequency of the term in Google n-gram corpus, presence of the term in Wordnet
UMLS derived features	Number of semantic types, synonyms, CUIs identified for the term;if the term is present in CHV; if the entire term has a CUI; if the semantic type of the term is one of the 16 semantic types from [24].

You were admitted for <u>coronary artery disease</u>. We provided treatment for anxiety with <u>calms</u> and alleviating anxiety. As a result, level of anxiety has improved slightly. We provided treatment for <u>self-bathing/hygiene deficit</u> with assistance in taking care[...] We treated <u>risk for infection</u> related to joint tuberculosis with <u>infection protection</u>. As a result, risk related to infectious process has improved slightly. FOLLOW-UP: follow up with CT sx in 1 week.

Figure 2. Part of the summary generated for Patient 156.

medical concepts that are present in the *hospital course* section of the doctor's discharge notes. MedLEE maps the entities to concepts in the Unified Medical Language System (UMLS) [23], which consists of more than 5 million concepts. The knowledge sources provided by UMLS allow us to query about different concepts, their unique id called Concept Unique Identifier (CUI) (e.g. CUI for the concept *coronary artery disease* is C0010054), and include definitions, along with the relationships between concepts. The NANDA, NIC, and NOC terms from the nursing POC are already included in the UMLS and also have corresponding CUIs. To generate the summary, for each patient, we build a concept graph by querying UMLS for CUIs that are related to the CUIs extracted from the doctors note and nursing POC. From the concept graph, we select only those CUIs that either belong to the original documents or are required to connect a pair of physician and nursing concept that would otherwise remain unconnected. These concepts are candidates for inclusion in our summary and are first passed through the Simplification module.

3.2. Simplification module

This module is responsible for identifying the health concepts that are difficult to understand and need further explanation. It consists of two components: a) a metric for identifying *Simple* and *Complex* terms, and b) a module for providing

definition to *Complex* terms. Unlike the existing methods that focus only on simplifying a single word, our metric works for concepts that consist of a single word like *arrhythmia* or multiple words like *coronary artery disease*. It extracts several features (see Table 1 for the complete list) from the concepts and uses a combination of clustering and linear regression approach for distinguishing between *Simple* and *Complex* concepts. If a concept is identified as being *Complex*, it is sent to the definition extractor and ranker unit (see Figure 1). This unit extracts definitions from three different knowledge sources: Wikipedia[2], Wordnet, and UMLS and performs the following steps: a) extracts all the health concepts present in the definitions by mapping them to UMLS, b) uses our metric for determining the score for each concept, c) adds the scores of the concepts in each definition and chooses the definition with the least score.

For instance, given a concept *coronary artery disease* which is identified as being *Complex* by our metric, the definition extractor and ranker module chose the definition from UMLS as being the simplest and hence the first occurrence of the term *coronary artery disease* in our summary will have the definition from UMLS attached to it. These concepts (with or without definition appended to them), along with relevant verbs are supplied as features to the phrasal operations provided by the SimpleNLG API [26], which then produces grammatically correct text. Figure 2 shows a part

---

[2]Wikipedia is also used by health care practitioners for getting familiar with the basics of some health issues [25].

of the summary[3] generated for Patient 156. All the terms that are underlined were identified as being *Complex* by the Simplification module and have a definition appended to them. This summary is further enriched during the personalization process.

4. Personalization

Personalization is the process of customizing a content based on the preferences and interests of different individuals. A survey of US Health Consumers in 2012 showed that a majority of the respondents were interested in creating plans that were specific to their needs rather than relying on pre-defined options [27]. Studies have shown that instead of than just narrowly focusing on disease and illness, effective health systems should also address other issues like the quality of the patient's life and their health and wellness [28]. Motivated by these studies, our personalization algorithm takes a wide range of aspects into consideration.

4.1. Features guiding personalization

Our personalized summary generation system takes the patient's ability to understand medical terms, their engagement in taking care of themselves, and their concerns and strengths into account. In order to identify the common issues and concerns, we conducted interviews with 26 patients who were hospitalized with heart issues. These interviews are open ended and the patients share their experiences since they were first diagnosed. The common topics that emerged from this analysis as well as other findings from the interviews are used for developing questions (see points A and B below) that will elicit a patient's strengths, concerns, and preferences in real time. We use an existing metric for determining the health literacy of the patient (see point C below) and a questionnaire to determine how engaged they are in taking care of their own health (see point

---

[3]Our algorithm did not find an explanation for the abbreviation "sx". The human-authored version of this summary also provides the same follow-up information. Depending upon the performance of the information extraction tool, some infelicitous concepts (like "calms" instead of "calming") appear in the summary.

D below). Hence, when patients are discharged, they will give the health literacy test, fill out the questionnaire that will determine their level of engagement, and answer the questions regarding their concerns and strengths (that were derived from the patient interviews).

A. Strengths/concerns of the patient: We are engaged in an ongoing annotation of the interviews and have coded 9 interviews so far. The coding categories are derived from our discussion with our collaborators in nursing and patient education. Through our preliminary analysis, we found several categories that most of the patients frequently mention, which are represented by the following four parameters and their possible values: 1) *priorities in life* (family/ friends/ health/ career/ employment/ finance/ religion), 2) *changes in life because of the health issue* (lifestyle/ diet/ mobility/ physical activity/ daily routine), 3) *means of support/ sources of strength or courage* (desire to get back to normal life/ family/ friends/ support groups/ caretaker), and 4) *ability to cope with health issues* (great/ able to manage it so far/ could have been better/ too bad). These parameters and their possible values will be presented as multiple choice questions to patients while they are getting discharged. Based on their response, personalized sentences will be generated and included in the summary. The possible values for these parameters are nowhere comprehensive and we will constantly update them as we code more interviews.

B. Patient's familiarity with the health issue: Based on our observations from the patient interviews (described above), we found that patients who have been suffering from the health issue for some time now or have someone in the family who already had the same issue before, seemed to be more familiar with the basic disease-specific terminologies. Hence, we introduce two parameters: a) *number of years since first diagnosis* with value *<3* or ≥ 3, where 3 is an arbitrary threshold that appears consistent with our observations from interviews, and we will verify in the future, and b) *history of the health issue in the family* with a *yes* or *no* value.

C. Health literacy: Health literacy represents

an individual's ability to process and understand basic health information and services needed to make appropriate health decisions [29]. We use a 66-itemed word recognition and pronunciation test called the REALM test, which depending upon how correctly a participant pronounces the words in the list, provides a score. This score tells us about which grade level the health literacy of the patient belongs to. We consider the following mapping between the health literacy level and our parameter *literacy score*: third grade or below=1, fourth to eighth grade=2, and high school level=3.

D. Patient engagement: The Patient Activation Measure (PAM) is a metric that quantifies what it means for a patient to be "engaged" in taking care of their health [30]. PAM consists of 13 questions that can be grouped under four developmental activation stages. Patients with a high PAM score follow their prescribed treatment and engage themselves in self-monitoring activities at home [31]. Low activated patients are poor self-managers and account for 75% of health care costs as well as 80% to 100% of readmits in care transition programs [32]. The PAM score can take a value of 1, 2, 3, or 4 depending upon the activation stage of the patient. We consider PAM scores of 1 and 2 as *low PAM* and 3 and 4 as *high PAM*. Apart from the PAM score, since three questions (Q4, Q8, and Q9) focus on capturing the patient's understanding of the health problem, treatments, and prescribed medications, we average their scores and store it in a parameter called *self-assessment*.

4.2. Personalization rules

Once the values for all the factors that are mentioned in Section 4.1 are obtained, we use a rule-based approach for making decisions on the content that should be included in the summaries. We also introduce a parameter called *overall familiarity*, whose value depends upon four other parameters: *literacy score*, *self-assessment* score, *number of years since first diagnosis*, and *history of the health issue in the family*. An aggregation function takes these 4 constituent parameters and structures them into a one level

Given: L= *literacy score*,
  Y= *number of years since first diagnosis*,
  H= *history of the health issue in the family*,
  Q= *self-assessment* score
Determine: *overall familiarity* (F)
If (L=1): If (($Y \geq 3$ and H=1) or ($Y \geq 3$ and $Q \geq 3$) or (H=1 and $Q \geq 3$))): F=2
  Else: F=1
Else if (L=2): If ($Y \geq 3$ and H=1 and $Q \geq 3$): F=3
  Else if (($Y \geq 3$ and H=1) or ($Y \geq 3$ and $Q \geq 3$) or (H=1 and $Q \geq 3$))): F=2
  Else: F=1
Else if (L=3): If (($Y \geq 3$ and H=1) or ($Y \geq 3$ and $Q \geq 3$) or (H=1 and $Q \geq 3$))):F=3
  Else: F=2

Figure 3. Rules for determining the value of overall familiarity parameter.

tree, whose structure and weights on branches are user-independent. Since we have given more weightage to *literacy score*, and the remaining 3 parameters have equal weights, there are three possible values for *overall familiarity* (as is the case for *literacy score*). Figure 3 shows our rules for determining the value of *overall familiarity*. Hence, *overall familiarity* and PAM score are two main factors that are responsible for personalizing the summaries that are generated by our system.

Similar to Ardissono et al. [33], we use a modular, compositional approach in which our two factors- *overall familiarity* and PAM score are managed separately. While *overall familiarity* is responsible for determining the level of detail and amount of information that should be included in the summary, PAM score uses the guidelines that are provided by the developers of PAM on how a patient with a particular score should be treated [30]. We provide more empathy to patients that have low PAM score, while we encourage the ones who have high PAM score to focus on maintaining their behavior. Our algorithm for personalized summary generation is shown in Figure 4 and two versions of summaries created for Patient 156 are shown in Figure 5 and Figure 6. As specified in the algorithm, the first paragraphs of both the versions explain the main health issue of the patient. Next, the low *overall familiarity* version (Figure 5) only mentions the other diagnoses (in the second paragraph), while the high *overall*

Given: M= Simplification module from Figure 1
For each patient P,
(A) Initialize:
NC=Nursing concepts (from P's plan of care)
PC= Physician concepts (from P's discharge note)
CG= Concept graph that maps the relation between all
    of the concepts in NC and PC (see Section 3.1)
Pam= Score from the PAM questionnaire (high/low)
Familiarity= *overall familiarity*
S1−S4= Response to the 4 multiple-choice questions
    from point A in Section 4.1
(B) Summary generation module:
1. i) Extract the main diagnosis from P's plan of care
  and pass it through M
  ii) Construct a sentence specifying the diagnosis. Also
  provide a brief explanation of the health issue.
2. If Familiarity=1:
  i) Extract all the other diagnoses from NC
  ii) Pass each diagnosis through M
  iii) Construct one or more brief sentences depending
  upon the number of diagnosis
  Else if Familiarity=2:
  i) Extract all the diagnoses and corresponding
  outcomes after treatment from NC
  ii) Pass each diagnosis and outcome through M
  iii) Construct sentences specifying the diagnoses and
  outcomes in separate paragraphs
  Else if Familiarity=3:
  i) Extract all the diagnoses from NC.
  ii) Extract the concepts from CG that are related to
  each diagnosis within a distance of 2
  iii) Extract all the interventions and outcomes
  corresponding to each diagnosis from NC
  iv) Pass each of the diagnosis, intervention and
  outcome through M
  v) Construct sentences specifying each diagnosis,
  additional concepts from CG, treatment, and outcome
3.Vary the amount of empathetic and encouraging
phrases in all the sentences depending upon whether
Pam=low/high.
4. Construct sentences from S1, S2, S3, S4
5. Extract and append follow-up information from P's
discharge note
6. If Familiarity=3, include a link that provides
additional information about the main diagnosis

Figure 4. Algorithm for generating personalized summary for a patient.

familiarity version (Figure 6) includes additional details like the diagnoses, interventions that were done to treat the issues, and the outcomes (in the second and third paragraph). Both the versions include the follow-up information, while only the high *overall familiarity* version includes a link to an additional source that the patient can refer to for learning more about the health issue. Similarly, for patients with low PAM score (Figure 5), we include more empathy and encouragement with sentences like *"Dealing with this issue must have been tough for you, we hope that you are feeling much better now."* and *"It is great that you are taking these important steps to improve your health"*. For patients with high PAM score, we motivate them to continue taking care of their health with sentences like *"We appreciate your efforts"* and *"Keep up the good work"*. The phrases that have been used for expressing empathy and encouragement, and the statements for reinforcing patient participation have been derived from the literature on physician-patient and nurse-patient communication [34, 35], as well as some online sources.[4,5]

5. Evaluation and Results

5.1. Coverage of medical terminologies

Since we summarize information from the discharge note and nursing POCs, we want an estimate of the coverage of "important" information from both the documents in our summaries. For this purpose, we asked a nursing student to read the physician and nursing notes for 35 patients and generate hand-written summaries. A doctor and a nurse went through 5 of the hand-written summaries and highlighted all the important concepts[6]. We compared the highlighted concepts in each hand-written summary with the concepts present in the corresponding automatically generated summary. On average,

---

[4]www.thedoctors.com/KnowledgeCenter/PatientSafety/ Appendix-2-Examples-Empathetic-Statements-to-Use
[5]www.kevinmd.com/empathy-patient-interactions.html
[6]While it would be better to have more than 5 summaries, our annotators in this case are practicing healthcare providers, whose availability is next to zero

Dear patient 156, we are sorry to know that you were admitted for coronary artery disease. Coronary artery disease is caused by impedance or blockage of one or more arteries that supply blood to the heart, usually due to atherosclerosis (hardening of the arteries). Dealing with this issue must have been tough for you, we hope you are feeling much better now.

During your hospitalization, you were monitored for chances of infection, problem in physical mobility and skin integrity. We also worked to improve self-bathing/hygiene deficit and anxiety level.

We can understand that you have to make changes in your way of living, diet and physical activity as a result of your health condition. It is great that you are taking these important steps to improve your health. Being committed to solving this problem is so important. We are very glad to know that you have sources to support you. We hope you feel better so that you can spend time with your family and friends and get more involved in religious activities.

FOLLOW-UP: follow up with CT sx in 1 week.

Figure 5. Low PAM score and low overall familiarity version of the summary for Pat. 156

our automatically generated summaries contain 80% of the concepts that have been highlighted as important by the doctor or nurse and 70% of the concepts from the handwritten summaries.

5.2. Interpretation of feature values

We created an online survey in Qualtrics and hosted it in Amazon Mechanical Turk (MTurk). Before running the study with MTurk, we piloted the summaries on three patient advisors (i.e. individuals who are patients themselves and are involved in research by providing their opinions) who provided feedback on various aspects of our personalized summaries. We then recruited 150 participants that are 18 years or older and are not employed in the healthcare industry. Our survey consists of 18 multiple choice questions (with no scaled-response questions), most of which present different situations to the user and ask them to choose an appropriate sample of text that a person facing the situation would most likely prefer. Since the questions do not reveal anything about the intended recipient of each sample, we can make

Dear patient 156, we are sorry to know that you were admitted for coronary artery disease. Coronary artery disease is caused by impedance or blockage of one or more arteries that supply blood to the heart, usually due to atherosclerosis (hardening of the arteries).

During your hospitalization, you were monitored for chances of infection, problem in physical mobility and skin integrity. We provided treatment for anxiety with calms and alleviating anxiety. We worked to improve self-bathing/hygiene deficit with assistance in taking care of oneself. We treated impaired mobility by engaging in light activities like walking, standing, etc. and giving encouragement to exercise. We worked to improve broken skin with wound care. We treated risk for infection related to joint tuberculosis with infection protection.

As a result of these interventions, ability to care for your hygiene, level of anxiety, and mobility have improved slightly. Risk related to infectious process and status of the skin have also improved slightly.

We appreciate your efforts in making changes in your way of living, diet and physical activity for maintaining your health. Keep up the good work. We are very glad to know that you have sources to support you. We hope you feel better so that you can spend time with your family and friends and get more involved in religious activities.

FOLLOW-UP: follow up with CT sx in 1 week.

For more information on coronary artery disease, please refer to the following website: https://www.mayoclinic.org/diseases-conditions/coronary-artery-disease/symptoms-causes/syc-20350613

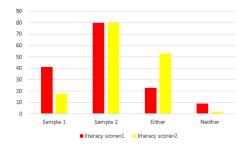**Figure 6. High PAM score and high overall familiarity version of the summary for Pat. 156.**

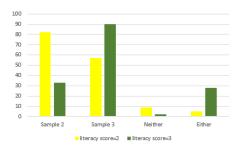**Figure 7. Samples chosen by participants for people with overall familiarity score of 1 or 2**

Figure 8. Samples chosen by participants for people with overall familiarity score of 2 or 3

Figure 9. Response on how patients with different PAM score can be treated

comparison between how the participant and the algorithm interpret the same situation.

Identifying the summary created for a patient with a certain *literacy score*: We created 3 samples of summaries (Sample1, Sample2, Sample3) with increasing amount of details and asked participants to choose the sample they would give to a patient with certain *literacy score*. For our example, Figure 5 without the third paragraph would count as Sample1 and Figure 6 without the fourth and sixth paragraph would count as Sample 3. Sample2 would add the third paragraph from Figure 6 to Sample1. The questions are framed in such a way that the participant needs to compare the given 2 versions (Sample1 vs Sample2, and Sample2 vs Sample3) and decide on the samples they would give to patients with lower and higher *literacy score*. We had expected that for Sample1 vs Sample2 (see Figure 7), more participants would choose Sample 1 for patients with lower *literacy score* because it has comparatively fewer medical details. Instead, participants preferred Sample2 for both the cases. Most of the participants who chose Sample2 for patients with lower *literacy score* said that "since Sample2 tells that certain aspects of the health of the patient have improved, it gives them hope", which reflects the opinion of the patient advisors as well. In case of Sample2 vs Sample3 (see Figure 8), most of the participants thought that Sample3 should be given to a patient with higher *literacy score*, as was expected.

Dealing with patients having *low PAM* and *high PAM*: Following the instructions provided by the developers of PAM, our algorithm focuses on
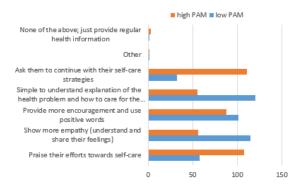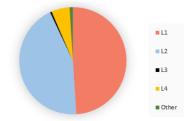
providing more empathy and encouragement for patients that have low PAM score. For the patients with high PAM score, we praise their efforts and encourage them to continue taking care of their health. As seen in Figure 9, these decisions were found to be consistent with the opinion of most of the participants of our survey.

We also found that more than 74% of the participants were able to correctly identify all the kinds of information (e.g. description of health issue, outcomes) that are present in different samples of our summaries. Their reasoning behind choosing a particular sample of summary for a patient with certain *literacy score* also matched the factors that were taken into consideration by our algorithm while generating the summary. Figure 10 shows the breakdown of the survey responses to question *"Do you have any comments on the samples of summaries shown in the survey?"*

6. Discussion and Conclusions

In this paper, we described our efforts on generating personalized hospital-stay summaries. These summaries are not intended to replace the materials that are currently provided to patients during discharge and should be considered as a supplementary material that can enhance the patient's ability to understand their health issues, and the next steps that they need to undertake. The evaluations indicate that our computer generated summaries have a good coverage of the health

L1: These summaries seem to be easier to read and understand as compared to the documents given in hospital
L2: If I was a patient; I would like to get such summaries when I am discharged
L3: I don't think such summaries can be helpful. I wouldn't read them.
L4: They are of reasonable length

**Figure 10. Breakdown of the participants'
feedback on our summaries**

concepts that are present in the hand-written summaries. Regarding the factors that guide our personalization process, no existing literature discusses how the different factors that contribute to a patient's familiarity with a health issue should be combined. We also cannot follow the existing approaches like that of Mairesse and Walker [10] where people are asked to rank features of restaurants based on their preference because it is impossible for an individual to determine which aspect of their knowledge about a health issue is more significant. Hence, we opted with providing user-independent weights. Similarly, apart from the instructions provided by the developers of PAM, there is no existing literature on the kind of phrases and sentences that should be used for communicating with patients that are in different stages of activation. Hence, we collected samples from working nursing professionals and from the literature on doctor-patient and nurse-patient communication. However, in order to make the content less repetitive, we need to find/develop more sources of such information. The feedback that we collected the evaluations show that even small steps like altering the level of details and acknowledging the concerns of the patients in the summaries can improve their perception towards health documents.

Our future plans include refining our inventory of patient concerns and strengths by conducting and coding more interviews. We

will improve our personalization algorithm by including more features and will fine-tune the system-assigned thresholds and weights. We plan to evaluate the usability and effectiveness of our summaries with real patients and assess whether their understanding of their health issues has improved after reading the automatically generated summaries.

References

[1] CDC, "Morbidity and mortality weekly report." Centers for Disease Control , https://www.cdc.gov/mmwr/index.html, 2016. Last accessed on 3/9/2018.

[2] G. K. Berland, M. N. Elliott, L. S. Morales, J. I. Algazy, R. L. Kravitz, M. S. Broder, D. E. Kanouse, J. A. Muñoz, J.-A. Puyol, M. Lara, *et al.*, "Health information on the internet: accessibility, quality, and readability in english and spanish," *Jama*, vol. 285, no. 20, pp. 2612–2621, 2001.

[3] A. F. Hernandez, M. A. Greiner, G. C. Fonarow, B. G. Hammill, P. A. Heidenreich, C. W. Yancy, E. D. Peterson, and L. H. Curtis, "Relationship between early physician follow-up and 30-day readmission among medicare beneficiaries hospitalized for heart failure," *Jama*, vol. 303, no. 17, pp. 1716–1722, 2010.

[4] C. K. McIlvennan, Z. J. Eapen, and L. A. Allen, "Hospital readmissions reduction program," *Circulation*, vol. 131, no. 20, pp. 1796–1803, 2015.

[5] U. D. of Health, H. Services, *et al.*, "Plain language: A promising strategy for clearly communicating health information and improving health literacy," *Retrieve from http://www. health. gov/communication/literacy/plainlanguage*, 2009.

[6] A. D. Boyd, K. D. Lopez, C. Lugaresi, T. Macieira, V. Sousa, S. Acharya, A. Balasubramanian, K. Roussi, G. M. Keenan, Y. A. Lussier, *et al.*, "Physician nurse care: A new use of umls to measure professional contribution: Are we talking about the same patient a new graph matching algorithm?," *International Journal of Medical Informatics*, 2018.

[7] C. H. Cain, E. Neuwirth, J. Bellows, C. Zuber, and J. Green, "Patient experiences of transitioning from hospital to home: an ethnographic quality improvement project," *Journal of Hospital Medicine*, vol. 7, no. 5, pp. 382–387, 2012.

[8] C. L. Paris, "Tailoring object descriptions to a user's level of expertise," *Computational Linguistics*, vol. 14, no. 3, pp. 64–78, 1988.

[9] I. Macdonald and A. Siddharthan, "Summarising news stories for children," in *Proceedings of the*

9th International Natural Language Generation conference, pp. 1–10, 2016.

[10] F. Mairesse and M. A. Walker, "Controlling user perceptions of linguistic style: Trainable generation of personality traits," *Computational Linguistics*, vol. 37, no. 3, pp. 455–488, 2011.

[11] A. J. Gill, C. Brockmann, and J. Oberlander, "Perceptions of alignment and personality in generated dialogue," in *Proceedings of the Seventh International Natural Language Generation Conference*, pp. 40–48, Association for Computational Linguistics, 2012.

[12] H. B. Jimison, L. M. Fagan, R. Shachter, and E. H. Shortliffe, "Patient-specific explanation in models of chronic disease," *Artificial Intelligence in Medicine*, vol. 4, no. 3, pp. 191–205, 1992.

[13] C. DiMarco, G. Hirst, L. Wanner, and J. Wilkinson, "Healthdoc: Customizing patient information and health education by medical condition and personal characteristics," in *Workshop on Artificial Intelligence in Patient Education*, 1995.

[14] S. Mahamood and E. Reiter, "Generating affective natural language for parents of neonatal infants," in *Proceedings of the 13th European Workshop on Natural Language Generation*, pp. 12–21, Association for Computational Linguistics, 2011.

[15] E. Erriquez and F. Grasso, "Generation of personalised advisory messages: an ontology based approach," in *Computer-Based Medical Systems, 2008. CBMS'08. 21st IEEE International Symposium on*, pp. 437–442, IEEE, 2008.

[16] G. M. Keenan, J. R. Stocker, A. T. Geo-Thomas, N. R. Soparkar, V. H. Barkauskas, and J. L. Lee, "The hands project: studying and refining the automated collection of a cross-setting clinical data set," *Computers Informatics Nursing*, vol. 20, 2002.

[17] T. H. Herdman, *Nursing Diagnoses 2012-14 : Definitions and Classification*. John Wiley & Sons, 2011.

[18] C. Strandell, "Nursing outcomes classification," *Rehabilitation Nursing*, vol. 25, no. 6, p. 236, 2000.

[19] H. K. Butcher, G. M. Bulechek, J. M. M. Dochterman, and C. Wagner, *Nursing interventions classification*. Elsevier Health Sciences, 2013.

[20] B. Di Eugenio, A. D. Boyd, C. Lugaresi, A. Balasubramanian, G. M. Keenan, M. Burton, T. G. R. Macieira, K. D. Lopez, C. Friedman, J. Li, *et al.*, "Patientnarr: Towards generating patient-centric summaries of hospital stays," *INLG 2014*, p. 6, 2014.

[21] S. Acharya, B. Di Eugenio, A. D. Boyd, K. D. Lopez, R. Cameron, and G. M. Keenan, "Generating summaries of hospitalizations: A new metric to assess the complexity of medical terms and their definitions," in *The 9th International Natural Language Generation conference*, 2016.

[22] C. Friedman, L. Shagina, Y. Lussier, and G. Hripcsak, "Automated encoding of clinical documents based on natural language processing," *Journal of the American Medical Informatics Association*, vol. 11, no. 5, pp. 392–402, 2004.

[23] NIH, "Umls quick start guide." Unified Medical Language System, https://www.nlm.nih.gov/research/umls/quickstart.html, 2011. Last accessed on 10/10/2016.

[24] B. P. Ramesh, T. K. Houston, C. Brandt, H. Fang, and H. Yu, "Improving patients' electronic health record comprehension with noteaid.," in *MedInfo*, pp. 714–718, 2013.

[25] B. Hughes, I. Joshi, H. Lemonde, and J. Wareham, "Junior physician's use of web 2.0 for information seeking and medical education: a qualitative study," *International journal of medical informatics*, vol. 78, pp. 645–655, 2009.

[26] A. Gatt and E. Reiter, "Simplenlg: A realisation engine for practical applications," in *Proceedings of the 12th European Workshop on Natural Language Generation*, pp. 90–93, Association for Computational Linguistics, 2009.

[27] P. H. Keckley and S. Coughlin, "Deloitte 2012 survey of us health care consumers," *Washington: Deloitte Development LLC*, 2012.

[28] A. Snowdon, C. Alessi, and K. Schnarr, *" It's All About Me": The Personalization of Health Systems*. International Centre for Health Innovation, Western University, 2014.

[29] D. A. Kindig, A. M. Panzer, L. Nielsen-Bohlman, *et al.*, *Health literacy: a prescription to end confusion*. National Academies Press, 2004.

[30] J. H. Hibbard, E. R. Mahoney, J. Stockard, and M. Tusler, "Development and testing of a short form of the patient activation measure," *Health services research*, pp. 1918–1930, 2005.

[31] K. Lorig, P. L. Ritter, D. D. Laurent, K. Plant, M. Green, V. B. B. Jernigan, and S. Case, "Online diabetes self-management program a randomized study," *Diabetes care*, pp. 1275–1281, 2010.

[32] InsigniaHealth, "Patient activation measure." Patient Activation Measure, https://www.insigniahealth.com/products/pam-survey, 2018. Last accessed on 2/18/2018.

[33] L. Ardissono, L. Console, and I. Torre, "An adaptive system for the personalized access to news," *Ai Communications*, vol. 14, 2001.

[34] E. C. Arnold and K. U. Boggs, *Interpersonal Relationships:Professional Communication Skills for Nurses*. Elsevier Health Sciences, 2015.

[35] E. J. Cassell, "Talking with patients: The theory of doctor-patient communication.(vol. 1)," 1985.