Know thy eHealth user: Development of biopsychosocial personas from a study of older adults with heart failure

Richard J. Holden, PhD[1,2,*], **Anand Kulanthaivel, MIS**[1], **Saptarshi Purkayastha, PhD**[1], **Kathryn M. Goggins, MPH**[3,4], and **Sunil Kripalani, MD, MSc**[3,4,5]

[1]Department of BioHealth Informatics, Indiana University School of Informatics and Computing – Indianapolis, IN, USA

[2]Indiana University Center for Aging Research (IUCAR), Regenstrief Institute, Inc. – Indianapolis, IN, USA

[3]Center for Effective Health Communication, Vanderbilt University Medical Center – Nashville, TN, USA

[4]Center for Clinical Quality and Implementation Research, Vanderbilt University Medical Center – Nashville, TN, USA

[5]Department of Medicine, Vanderbilt University School of Medicine – Nashville, TN, USA

Abstract

Background—Personas are a canonical user-centered design method increasingly used in health informatics research. Personas—empirically-derived user archetypes—can be used by eHealth designers to gain a robust understanding of their target end users such as patients.

Objective—To develop biopsychosocial personas of older patients with heart failure using quantitative analysis of survey data.

Method—Data were collected using standardized surveys and medical record abstraction from 32 older adults with heart failure recently hospitalized for acute heart failure exacerbation. Hierarchical cluster analysis was performed on a final dataset of n=30. Nonparametric analyses were used to identify differences between clusters on 30 clustering variables and seven outcome variables.

*Corresponding Author: Richard J. Holden, PhD, Walker Plaza – WK317, 719 Indiana Avenue, Indianapolis, IN 46202, USA. (ph) 1-317-278-5323, rjholden@iupui.edu.

Results—Six clusters were produced, ranging in size from two to eight patients per cluster. Clusters differed significantly on these biopsychosocial domains and subdomains: demographics (age, sex); medical status (comorbid diabetes); functional status (exhaustion, household work ability, hygiene care ability, physical ability); psychological status (depression, health literacy, numeracy); technology (internet availability); healthcare system (visit by home healthcare, trust in providers); social context (informal caregiver support, cohabitation, marital status); and economic context (employment status). Tabular and narrative persona descriptions provide an easy reference guide for informatics designers.

Discussion—Personas development using approaches such as clustering of structured survey data is an important tool for health informatics professionals. We describe insights from our study with heart failure patients, then recommended a generic ten-step personas development process. Methods strengths and limitations of the study and of personas development generally are discussed.

INTRODUCTION

User-centered design (UCD) is crucial to creating useful, usable, and satisfying health information technology (IT) application [1, 2]. The chief UCD principle is to base design on a strong understanding of the intended users, summarized by the dictum *know thy user*. Increasingly, leaders in eHealth and consumer health IT in particular have called for the application of this and other UCD principles [3, 4] in parallel with further development of UCD methods to accommodate the unique nature of eHealth [5, 6]. *Personas* is one UCD method worth adapting, implementing, and illustrating for the eHealth context. Personas are fictitious user archetypes based on real (qualitative, quantitative, or mixed) data used by IT designers to imagine the kind of users to be accommodated [7]. For example, a team designing a mobile application to help older adults manage medications would examine several older adult personas to achieve a sense of typical needs, range of medications taken, and variation in technological competency. A designer might attempt to accommodate the range or else might design alternatives or add-ons to help a specific type of user (e.g., smartphone novices). Personas are also used in expert evaluation of products, for designing or recruiting for usability testing, and for marketing and education during product deployment [7].

While personas are ubiquitous in UCD for IT design outside of healthcare, there are few published, detailed accounts of personas development for health IT. Personas are recommended as part of user-centered process for consumer health IT design and implementation, in particular, as a simple tool for communicating with the many stakeholders involved in health IT design, procurement, deployment, and management [8]. However, not only are there few examples of such work, but there are fewer studies using quantitative data for personas development [9–11]. Among those studies, none has considered the full range of biological, psychological, and social (i.e., biopsychosocial)

variables that characterize actual patient users. Therefore, we report on a study in which we used a systematic approach to patient user personas development, using quantitative cluster analysis on biopsychosocial survey data from older patients with heart failure.

Personas: a user-centered design tool

The international standard on UCD, ISO Standard 9241-210, dictates that design begins with an empirical definition of users and their context [12]. There is limited guidance on achieving this empiricism, but one popular general approach is developing personas based on data collected from the intended user population [13].

In introducing personas for IT development, Cooper [14] defined a persona as a "hypothetical archetype of actual users…defined with significant rigor and precision." Personas are usually generated from the study of a population and constitute a scientific model that makes sense of volumes of chaotic information on users and their goals. The most effective personas are empirical products, though these can be supplemented with knowledge from subject matter experts (e.g., physicians, researchers) [15]. There is no single accepted way to form the archetypes and map actual participants to the archetypes, but options include using qualitative thematic analysis, affinity diagrams, or factor analysis of quantitative data [13, 16]. It is often argued that personas can accelerate product design, enhance communication with customers, and contribute to post-launch interface development [7, 15]. Indeed, Miaskiewicz et al. [17] identify 22 specific benefits of personas, with leading benefits including focus on specific (customer) audiences, guiding design and prioritization decisions, and challenging designers' assumptions. Not surprisingly, personas are widely recommended (e.g., by usability.gov) and used by popular software developers such as Microsoft Corporation, despite questions about the rigor of personas methodology [18].

Personas for eHealth

Despite widespread use in other industries, personas are rarely reported in healthcare, clinical health IT, or consumer health IT literature [9]. One recent example is the "Voices of Veterans" project at the Veterans Affairs (VA) Center for Innovation, whose Human-Centered Design methodologies produced seven personas varying in demographics, education, occupation, needs, attitudes, medical benefits used, and personal background [19]. The project describes how designers could use the personas as well as future directions in personas development, such as personas evolution and refinement. To account for the dynamics of patient experience, as opposed to creating static patient profiles, Hall et al [20] develop journey maps of cancer patients to guide design. In a study of older Chinese adults with diabetes, LeRouge et al. [9] performed qualitative data collection (focus groups, observations, interviews) and qualitative analysis to create a number of participant profiles. The authors presented two illustrative personas, a rural 68 year-old married Chinese woman and an urban 63 year-old married Chinese woman, and described ways personas could be used for UCD projects on eHealth. Of interest, LeRouge et al.'s patient profiles and personas deliberately included variables such as social support network and attitudes towards providers in order to supplement traditional demographic and cognitive factors such as education, computer skills, and learning style. Valdez et al. [6] recently argued the

importance of taking this broader, *contextual* or *biopsychosocial* approach to understanding current or prospective consumer health IT users. The authors introduced a hypothetical persona based on an amalgam of patients from prior studies, "Brenda," a 48-year-old woman with diabetes, and the physiological, social-behavioral, and contextual (e.g., social, organizational, and environmental) factors shaping her everyday activity and technology use. While both LeRouge et al. and Valdez et al. urged attending to biopsychosocial factors during consumer health IT design, neither quantitatively assessed these factors. Other limitations of existing literature include inadequate description of the personas development method [21] or the actual personas yielded [22].

Biopsychosocial personas development for older patients with heart failure

In the present study, we used quantitative methods to develop personas inclusive of a range of biopsychosocial variables relevant for eHealth design. We chose heart failure (also known as chronic heart failure and congestive heart failure) as the illustrative patient user domain and used data from a survey of older patients with heart failure. Heart failure is a chronic, terminal illness especially common in older adults, with 4.6 million cases estimated among Americans aged   65. Similar population prevalence rates are found across North America, Europe, Asia, and Australia [23]. Heart failure is a common reason for hospitalization and rehospitalization [24], and a major target for disease management and readmission reduction efforts. It imposes both symptoms and self-care requirements that affect quality of life. Studies report several meaningful differences among patients with heart failure, including differing levels of knowledge and expertise, motivation and engagement, cognitive function, social and instrumental support, and living arrangements [6, 11, 25]. These and other factors can be combined to form individual patient profiles or archetypes of patients (i.e., personas), to be used by designers of eHealth and other interventions. Our present research objective was **to develop biopsychosocial personas of older patients with heart failure using quantitative analysis of survey data**.

METHOD

The study was a secondary analysis of survey data collected on a sample of 32 older adults with chronic heart failure who were recently hospitalized for acute heart failure or heart failure and myocardial infarction at an academic medical center in the Southeast US. The primary dataset for this analysis was collected through the Vanderbilt Inpatient Cohort Study (VICS), described in detail elsewhere [26]. The VICS survey was a standardized questionnaire administered by a researcher during hospitalization, and at approximately two, 30, and 90 days post-discharge. Additional data were collected through the Caring Hearts Study (see [27]) using semi-structured interviews and standardized self-administered surveys a mean of 57.5 days (SD=19.5) post-discharge. Lastly, electronic medical records were abstracted for diagnoses, CHF functional status and left ventricular ejection fraction (LVEF), medications, laboratory test results, hospitalizations, and death, some of which were used in the analyses. Participants were patients aged 65 or older living within a 300-mile radius covering multiple regions of two states.

Procedure

The VICS survey [26] and Caring Hearts Study instruments [27] had over 500 items combined spanning multiple domains from biological and physiological, to psychological and behavioral, to social and economic. Variables from these instruments were selected for inclusion in personas development to cover the range of biopsychosocial domains deemed useful for eHealth design. The domains were selected in advance based on patient-centered sociotechnical systems models such those in Valdez et al [6], Holden et al [27–30], and reports of the National Research Council [31] and World Health Organization [32]. In the next step, two of the authors (RH and AK) performed a consensus-based card sort to group survey items or scales and medical record data into domains, yielding 48 subdomains. After inspection of data for each of the 48 subdomains, a subset of 37 were retained because they had data from at least 30 participants and values were not identical. Seven of the subdomains were considered outcome variables and were not used in personas development. Instead, personas were compared on these outcomes. Thus, the factors in the personas cluster analysis were 30 subdomains across these domains: demographic, medical, functional, psychological, technological, behavioral, health system, social context, and economic context (Table 1). (For all 48 original domains and measurement details, see Appendix A.)

Data analysis

Prior to analysis, two of the 32 participants were eliminated due to item non-response (< 60% of questions answered), yielding a final analyzed set of N=30. We then log-transformed variables with power-law or leptokurtotic-type distributions (standard error of distribution skew   3.0). The following scales were calculated from constituent survey items: Subjective Numeracy Scale [40], Wake Forest Physician Trust Scale [49, 50], Perceived Health Competence Scale [37], Brief Resilient Coping Scale [41], Adherence to Refills and Medications Scale [48], Brief Health Literacy Scale [36], Vulnerable Elders Survey [34], and Patient Health Questionnaire [35].

Clustering was then performed using hierarchical cluster analysis (HCA) and Ward's method. HCA compares each observation (patient) to all others and places observations into clusters based on similarities. Ward's method [51] uses an agglomerative approach wherein each observation is assigned its own cluster and then cluster pairs are progressively grouped based on similarity until either a distance or number of cluster criterion is achieved. The end product of an HCA is a dendrogram or branch diagram of clusters and observations, with similarity represented by physical distance between branches. HCA differs from clustering techniques like k-means clustering, which pre-selects a number of clusters (k), then assigns observations to each based on proximity between the observation and the cluster mean. HCA is considered an appropriate technique for clustering in dimensional, low sample size (HDLSS) data and of particular present relevance has been demonstrated to be valid in sample sizes of   30 [52–54]. HDLSS data sets are ones that have more columns (variables) and fewer rows (respondents) of data, like our survey data. Multivariate cluster analysis of our dataset was performed in MINITAB 17 with Ward's algorithm and a stopping rule of six clusters. The stopping rule of six clusters was selected to balance the number of members in the cluster in relation to the small overall sample size, per common practice [55, 56].

After obtaining the clustering results, we tested the statistical significance of differences between the six identified clusters for each subdomain variable. Welch's Analysis of Variance (ANOVA) was performed on continuous variables and the nonparametric Wilcoxon-type Kruskal-Wallis (K-W) ANOVA on binomial variables, tests appropriate for overcoming issues with heteroscedasticity in smaller samples [57]. They were performed using SAS (SAS Inst., Cary, NC) and judged at the threshold $p < 0.10$.

Lastly, a persona was created for each cluster. Only those variables that differed significantly between clusters were included in the personas descriptions. The z-score was calculated for cluster values on each continuous variable, and the proportion calculated for each binomial variable. Based on this, we generated personas narratives and qualitative descriptors. For continuous variables, such as age, we used the cluster mean for that variable as the value for the persona (e.g., the persona's age). For categorical variables, we used the rule of thumb of a z-score < -0.40 or > 0.40 to identify the attributes salient to that cluster. For example, the z-score for health literacy was examined for each cluster, then: a) the persona with lowest z-score of -1.94 was labeled "least health literate;" b) the persona with the highest z-score of 0.86 was labeled "most health literate"; b) the personas with z-scores of < -0.40 (but not the lowest) were labeled as "less health literate" and those with z-scores of > 0.40 were labeled "more health literate"; and c) personas with z-scores between -0.40 and 0.40 were given no label on this variable because it was not salient for them. For binomial variables such as gender, the cluster's label was definitive ("male" or "female") if the cluster was homogeneous (all male or all female), "likely" if skewed towards one value (e.g., "likely female" if the cluster was mostly female), and "possibly" if proportionately balanced (e.g., "possibly male" if the cluster was evenly divided on gender). Once narrative labels were developed from cluster data, we performed several checks to ensure that the wording was reasonable given the quantitative scores, but did not refer back to original survey instruments to match our labels to descriptors or interpretations in the literature (e.g., cut-off scores for "adequate" vs. "inadequate" health literacy). When writing the final persona descriptions, we used factual language, as above, reflecting the z-score and proportion rules above.

Because of the focus on health IT, we performed an additional qualitative content analysis of each patient's semi-structured interview data to identify technology experience and perceptions. These technology attributes were not included in the cluster analysis but are used to supplement personas descriptions. Personas demographics were based on the modal demographics of the participant in that persona's cluster. Two of the personas, representing two of the more atypical clusters, were further transformed into a one-page persona document, as is typical in the practice of creating personas. The document provides easy to read narrative information on the persona and a stock photo with representative quotation for the purpose of realism.

RESULTS

Of the 30 patients included in analyses, 53% were male, 87% were White non-Hispanic, and the mean age was 72.13 (median = 69, SD = 6.54, range 65–84). Based on 2010 US Census data, 54% of 28 patients reporting an eligible zip code lived in an urban area, 21% in an area with rural population 20–50%, and 25% in an area with a rural population 51–100%. The

majority (63%) were married and 80% lived with someone else. Five patients (17%) were employed part- or full-time. Most patients (90%) reported Internet access but only 27% had registered for or to some extent used their health system's online patient portal.

Personas clusters

Cluster analysis yielded six distinct clusters ranging in size from two to eight patients, illustrated in the dendrogram in Figure 1. Of the 37 total variables used, 16 were significantly different between clusters at the selected p<0.10 threshold; 12 of these were significantly different at the p<0.05 threshold. The p-values for all variables are reported in Table 2.

Table 3 shows the demographics and other attributes of each of the six personas, one per cluster, A – "Angela" through F – "Frederick." Depicted attributes were variables that differed significantly between clusters or were significant demographic (race) or outcome (portal use) factors (Appendix B provides the corresponding Z-scores and binomial proportions used to produce Table 3). From the table of attributes, a persona narrative can be written about each, as follows.

"Angela" (N=4 in this cluster)—Angela is 70 years old, White, and lives with her husband. She is easily able to do chores around the house, and is not diabetic. Despite rating her ability to do physical activities as very high, her actual physical exercise is minimal. She rates her mood as rather depressed. She has Internet access at home and is the most likely to use the patient portal system.

"Barbara" (N=8 in this cluster)—See Table 4.

"Charlene" (N=6 in this cluster)—Charlene is 75 years old, White, widowed, lives alone, and works part time. She has type 2 diabetes. She was able to receive assistance in transition from her recent hospital discharge, and has been visited by a clinician recently. Despite having Internet access at home, she does not use the patient portal system. She is able to take care of her hygiene.

"David" (N=6 in this cluster)—David is 69, White, and lives with his wife. He has many favorable factors, including high health literacy, high functional ability, and having someone care for him in transition out of the hospital. He reports no problems with depression. At home, he has Internet access and has used the portal to renew medications.

"Earl" (N=4 in this cluster)—Earl is 71, White, and lives with his wife. He is able to care for his own hygiene; however, he reports poor abilities with respect to household chores and general physical activities. Earl likes to think in terms of numbers, but has trouble understanding medical documents. Despite having Internet access at home, he has never used his health system's patient portal.

"Frederick" (N=2 in this cluster)—See Table 4. Table 4 presents more detailed examples of the final two personas, "Barbara" and "Frederick," in a one-page graphic

format. These documents are often used by designers to imagine with some level of realism the kind of individuals for whom they are designing.

DISCUSSION

This study performed cluster analysis on survey data representing a variety of biopsychosocial variables to produce six personas of older adults hospitalized with decompensated heart failure. Individuals in each cluster had similar characteristics to one another but differed on multiple variables from individuals in other clusters. For each cluster, a persona could be created, highlighting variables known to vary between clusters, which may be important for design of eHealth and other interventions. For example, for the two personas Barbara and Frederick, designers might have distinct considerations:

- Barbara is notable for being a receptive user of eHealth who is already using her health system's patient portal. However, Barbara represents only 20% of the sample (6/30). A designer should understand that users like Barbara may be primed to use new eHealth but may have specific needs. For example, Barbara is diabetic, depressed, and takes many medications. She might benefit from an IT intervention that supports self-management and goal-setting for exercise, diet, and symptom monitoring. The system should engage and motivate Barbara, perhaps by allowing her clinicians to tailor content or send encouraging feedback [58, 59].

- Frederick is fairly unique, representing only two respondents (6.67%), but cannot be ignored by designers. Frederick does not have Internet access and is uncomfortable with new technology. This is what Adlin and Pruitt [7] call an "anti-persona," a type of person who might not use the IT being developed. However, when we examined qualitative data from the two patients in this cluster, we found both were willing to have a family member use technology on their behalf. This suggests opportunities to design informal caregiver-centered eHealth for Frederick and people like him [60].

The full set of personas and the expanded personas of Barbara and Frederick are particular to the analysis of data from older adults with heart failure but can be contrasted to other health personas, such as those from LeRouge and colleagues' [9] qualitative study of older adults with diabetes. By comparison, that study's persona characteristics were more subjective and imprecise (e.g., "outgoing and warmhearted") but also richer, describing personal history, familial relations, and hobbies. Both described biopsychosocial characteristics but LeRouge et al.'s personas included a qualitative analysis of information needs, providing more direct implications for eHealth design.

The use of biopsychosocial variables including elements of the person and their context is useful when considering chronic illness and heart failure in particular, given the multiple, cross-level factors affecting chronic illness clinical management and heart failure self-care [27–29]. For example, the importance of comorbid diabetes in our heart failure study suggests self-management technologies supporting self-management for both conditions. Another design implication is designing to support the range of literacy and numeracy found

across our six heart failure patient personas. A third example implication relates to differences between those who are discharged from the hospital with or without formal and informal caregiving support in the home: for those who do not have such support, technology could provide access to one's physically distributed social network or locally available social services. These and other biopsychosocial factors play a complicated role in determining the value of IT for older adults with heart failure and other conditions [29, 61], which means they must be considered by eHealth designers. The complex identities and contexts of patients and other health-seekers is one reason multidimensional personas and other UCD methods have been recommended in the domain of health and healthcare [6, 62].

Using personas development methods for eHealth research and practice

Tools such as personas may help designers to take a variety of potentially relevant factors into account in building eHealth systems and services such as smartphone apps, social media platforms, and telemedicine [63]. This study used standardized survey methods, which yielded data on many variables in a manner more efficient than extensive qualitative interviews or observations. The use of self-rather than researcher-administered surveys could even further increase the efficiency of data collection, generating much larger data sets than the one analyzed here. For example, one impressive personas-development study outside healthcare by researchers at Sun Microsystems analyzed data from over 1,300 survey respondents [13]. In healthcare practice, incorporating survey instruments of our length or longer is feasible in some cases, particularly in the context of funded academic research. However, it may not be deemed worth the additional burden placed on patients or staff. Incorporating rich survey assessments into a process such as hospital discharge would have to either serve multiple practical purposes besides personas generation or else administrators would have to agree on the value of personas in their operational efforts. The latter would require further evidence that personas can be translated into useful eHealth product requirements and design recommendations. Furthermore, studies may be needed showing that the design of eHealth using personas produces more effective, usable, or acceptable technologies compared to other methods. On the other hand, the increasing availability of data from multiple sources including the electronic health record (EHR), patient-reported data, and wearable devices [64], may make data-driven personas generation increasingly more practical.

The advantages of personas for design have been thoroughly described elsewhere [7, 17]. However, healthcare researchers and practitioners may be reluctant to use personas development methods for a variety of reasons. These might include the costs and effort of the method as described above. Some may perceive personas development as an industry approach without scientific basis, despite attempts to introduce theoretical and practical bases to personas methods [16, 65]. Perceptions of inadequate scientific basis may stem from disagreement in the literature about the definition of personas [65] and lacking consensus on personas-generating methods [66]. Some note that the specific methods used across various projects are often inadequately described [67] and commercial entities are reluctant to share their methods and personas with the public [20]. Further perceived weaknesses are that personas may not be verifiable [18] and can appear to be impersonal, misleading, or caricatures lacking nuance [67].

Although it is outside the scope of this paper to address each perceived weakness of personas, we believe a full reporting of one's personas development methods is at least one way to promote their use, while addressing criticism about lack of scientific rigor. Therefore, to facilitate others' evaluation, modification, and replication of our approach, Table 5 presents a ten-step process for personas development and use based on the approach we used.

Methods issues for generating personas

The present study applied quantitative methods for personas development and we note several strengths and limitations of this approach for developing eHealth. Advantages include efficiency in the collection and analysis of data and the ability to cluster individuals on multiple simultaneous variables. Further, the same instruments can be used across samples and to aggregate or compare between groups. For example, it would be possible to compare personas for hospitalized (decompensated) vs. non-hospitalized patients with heart failure or to create personas representing chronically ill older adults across disease conditions. Limitations of a quantitative approach are the lack of depth of understanding of not only each variable (e.g., the specific nature of available social support) but how these variables might relate to specific information needs that could be addressed by eHealth (e.g., the role eHealth could play in connecting patients to social support networks). Without deeper, more contextualized knowledge about the subdomain variables in this analysis, it would potentially be difficult for an eHealth designer to translate the clusters into specific meaningful designs. The resultant design considerations might therefore be generic, though nevertheless useful, for example, "take health literacy into account." Another limitation is that while quantitative cluster analysis may produce many clusters, eHealth designers may not be able to accommodate so many variants or user types; this may be addressed by clustering the clusters into broader categories.

Recent innovations in personas development research have included promising alternatives such as secondary analysis of video observations data [68] and use of objective telemetry and computer clicks [69]. These methods may reduce the subjectivity of self-report but introduce the potential for bias on the part of the analyst. An alternative approach is to use multiple semi-structured or ethnographic methods [70–72], as demonstrated in a recent comprehensive ethnographic study by Burrows and colleagues [73], who used a combination of semi-structured interviews, home tours, cultural probes, diaries, and focus groups to propose 15 attributes of home healthcare technology users. Less structured methods permit the capture of a broader range of attributes and deeper understanding of each, at the cost of higher subjectivity and greater data collection and analysis effort. In our recent design of a heart failure self-care health IT, we demonstrate the value of using largely qualitative data from a multiyear study but also identify the challenges of prolonged analysis and the difficulty of incorporating a rich and heterogeneous set of findings into a single design [74]. Other health researchers have demonstrated faster health-related persona generation techniques, for example a 90-minute rapid personas development method using a single facilitated session involving multiple individuals with deep knowledge of the patient population, but not using direct patient-generated data [75].

Limitations

This demonstration study had a small sample size for studies using cluster analysis. To compensate, we employed HCA for clustering and nonparametric tests to compare clusters. As mentioned earlier, HCA was an acceptable method to use because of the high proportion of variables to participants in our high-density dataset. Nevertheless, there were several drawbacks to the small sample size. First, it contributed to at least one very small and potentially unstable cluster (n=2), despite choosing a stopping rule of 6 clusters to avoid identifying a large number of small sized clusters. Second, the small sample size may have inflated Type II error during comparisons of clusters on subdomain variables, reducing the probability of discovering actual differences between clusters on variables of relevance to eHealth design. Although there is no agreement on rules of thumb for cluster analysis sample size, an increasing numbers of variables should be accompanied by increasing samples. An ideal study designed for personas development would have at least five participants per clustering variable, i.e., n=150 for a 30 variable analysis.

Our sample also had little variability with respect to certain factors such as race. Several variables selected for our study were excluded prior to analysis due to missing data or low variability, which could have been addressed during sampling and data collection. The use of an alpha threshold of 0.10 is acceptable for exploratory studies but increases the probability of Type I error. In the present analysis, this means clusters may not differ as much as reported. The lack of difference between clusters on several variables may limit the design insights usually sought from personas, which are often designed to highlight major differences between target end-users. An example of this in our study was internet availability, which in part led to the splitting into three clusters of otherwise relatively similar patients. We note that the most likely variables to differ between clusters were demographic factors, though variables representing multiple biopsychosocial domains differed between clusters and were used to construct personas. This study collected but only minimally incorporated qualitative data in its personas descriptions. Future research should attempt to combine quantitative and qualitative data to produce richer, contextualized descriptions of personas. This can help to empirically understand the nature of each persona, for example, explaining how Frederick's fatigue is the result of his physical therapy or that Barbara's multiple medications are related to comorbid heart failure, diabetes, and depression. An attempt could also be made to combine qualitative and quantitative data in generating the personas, using mixed-method analysis.

Other innovative future approaches could include: involving patients directly in persona creation and validation [76]; developing personas that represent dyads rather than individuals [77]; and updating personas over time to reflect personal or population trends such as increasing personal IT ownership [78], social media use [79], and Internet access [80] across demographic categories. An important future consideration is to broaden the sample beyond a single medical condition and to develop broadly relevant personas representing a variety of health needs, demographics, and biopsychosocial characteristics. Furthermore, future research should examine using personas methods for developing eHealth and medical devices for use by healthcare professionals [72, 81]. Additional

research must also demonstrate the practical value and feasibility of personas methods for actual eHealth development.

CONCLUSION

To our knowledge, this research represents a unique demonstration of quantitative personas-generation focused on older adults with chronic disease, in this case, heart failure. The personas-generation approach, described here and recommended for use by others, is a useful tool for designers of eHealth and other interventions for patients, clinicians, and other users. It may become more useful still with the increasing availability of data from and about patients. We urge others to use both the geriatric heart failure personas generated from this research and the methods described here to get to "know thy user."

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Holden, RJ., Voida, S., Savoy, A., Jones, JF., Kulanthaivel, A. Human Factors Engineering and Human–Computer Interaction: Supporting User Performance and Experience. In: Finnell, J., Dixon, BE., editors. Clinical Informatics Study Guide. Springer; New York: 2016. p. 287-307.

2. Middleton B, Bloomrose M, Dente MA, Hashmat B, Koppel R, Overhage JM, Payne TH, Rosenbloom ST, Weaver C, Zhang J. Enhancing patient safety and quality of care by improving the usability of electronic health record systems: Recommendations from AMIA. J Am Med Informatics Assoc. 2013; 20:e2–e8.

3. Zayas-Caban T, Dixon BE. Considerations for the design of safe and effective consumer health IT applications in the home. Qual Saf Health Care. 2010; 2010(Suppl 3):i61–i67.

4. Goldberg L, Lide B, Lowry S, Massett HA, O'Connel T, Preece J, Quesenbery W, Shneiderman B. Usability and accessibility in consumer health informatics: Current trends and future challenges. Am J Prev Med. 2011; 40:5187–5197.

5. Marquard J, Zayas-Caban T. Commercial off-the-shelf consumer health interventions: Recommendations for their design, evaluation, and redesign. J Am Med Informatics Assoc. 2012; 19:137–142.

6. Valdez RS, Holden RJ, Novak LL, Veinot TC. Transforming consumer health informatics through a patient work framework: Connecting patients to context. Journal of the American Medical Informatics Association. 2015; 22:2–10. [PubMed: 25125685]

7. Adlin, T., Pruitt, J. The Persona Lifecycle: Keeping People in Mind Throughout Product Design. Morgan Kaufmann, San Francisco: 2010.

8. Van Velsen L, Wentzel J, Van Gemert-Pijnen J. Designing eHealth that matters via a multidisciplinary requirements development approach. JMIR Research Protocols. 2013; 2 http://www.researchprotocols.org/2013/2011/e2021/.

9. LeRouge C, Ma J, Sneha S, Tolle K. User profiles and personas in the design and development of consumer health information technologies. Intl J Medical Informatics. 2013; 82:e251–e268.

10. Petersen, L., Bertelsen, P. Digital Healthcare Empowering Europeans. IOS Press; Amsterdam, The Netherlands: 2015. Citizen personas: Exploring challenges of citizen-centric eHealth; p. 582-586.

11. Vosbergen S, Mulder-Wiggers JM, Lacroix JP, Kemps HM, Kraaijenhagen RA, Jaspers MW, Peek N. Using personas to tailor educational messages to the preferences of coronary heart disease patients. Journal of Biomedical Informatics. 2015; 53:100–112. [PubMed: 25239261]

12. Jokala, T., Iivari, N., Matero, J., Minna, K. The standard of user-centered design and the standard definition of usability: Analyzing ISO 13407 against ISO 9241-11; Proc Latin American Conference on Human Computer Interaction (ACM); 2003. p. 53-60.

13. McGinn, J., Kotamraju, N. Data-driven persona development; Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08); 2008. p. 1521-1524.

14. Cooper, A. The Inmates Are Running The Asylum. Sams Publishing; Indianapolis, IN: 1999.

15. Cooper, A., Reimann, R., Cronin, D., Noessel, C., Csizmadi, J., LeMoine, D. About Face: The Essentials of Interaction Design. Wiley; San Francisco, CA: 2014.

16. Pruitt, J., Grudin, J. Personas: practice and theory; Proceedings of the 2003 conference on Designing for user experiences, ACM; 2003. p. 1-15.

17. Miaskiewicz T, Kozar K. Personas and user-centered design: How can personas benefit product design processes? Design Studies. 2011; 32:417–430.

18. Chapman C, Milham R. The persona's new clothes: Methodological and practical arguments against a popular method. Proceedings in Human Factors and Ergonomics. 2006:634–636.

19. VA Center for Innovation. Voices of Veterans: Introducing Personas to Better Understand Our Customers. U.S. Department of Veterans Affairs; 2014. https://www.vets.gov/playbook/downloads/Voices_Of_Veterans.pdf

20. Hall L, Kunz B, Davis E, Dawson R, Powers R. The Cancer Experience Map: An approach to including the patient voice in supportive care solutions. J Med Internet Res. 2015; 17 http://www.jmir.org/2015/2015/e2132/.

21. Serio CD, Hessing J, Reed B, Hess C, Reis J. The effect of online chronic disease personas on activation: within-subjects and between-groups analyses. JMIR Research Protocols. 2015; 4 http://www.researchprotocols.org/2015/2011/e2020/.

22. Kayser L, Kushniruk A, Osborne R, Norgaard O, Turner P. Enhancing the effectiveness of consumer-focused health information technology systems through eHealth literacy: A framework for understanding users' needs. JMIR Human Factors. 2015; 2 https://humanfactors.jmir.org/2015/2011/e2019/.

23. Ponikowski P, Anker SD, AlHabib KF, Cowie MR, Force TL, Hu S, Jaarsma T, Krum H, Rastogi V, Rohde LE, Samal UC, Shimokawa H, Budi Siswanto B, Sliwa K, Filippatos G. Heart failure: Preventing disease and death worldwide. ESC Heart Failure. 2014; 1:4–25. [PubMed: 28834669]

24. Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the Medicare fee-for-service program. New England J Med. 2009; 364:1418–1428.

25. Riegel B, Moser D, Anker S, Appel L, Dunbar S, Grady K, Gurvitz MZ, Havranek EP, Lee C, Lindenfeld J, Peterson P, Pressler S, Schocken D, Whellan D. State of the science: Promoting self-care in persons with heart failure: A scientific statement from the American Heart Association. Circulation. 2009; 120:1141–1163. [PubMed: 19720935]

26. Meyers G, Salanitro A, Wallston K, Cawthon C, Vasilevskis E, Goggins K, Davis C, Rothman R, Castel L, Donato K, Schnelle J, Bell S, JS S, Osborn C, Harrell F, Kripalani S. Determinants of health after hospital discharge: Rationale and design of the Vanderbilt Inpatient Cohort Study (VICS). BMC Health Serv Res. 2014; 14 https://bmchealthservres.biomedcentral.com/articles/10.1186/1472-6963-1114-1110.

27. Holden RJ, Schubert CC, Mickelson RS. The patient work system: An analysis of self-care performance barriers among elderly heart failure patients and their informal caregivers. Applied Ergonomics. 2015; 47:133–150. [PubMed: 25479983]

28. Holden RJ, Schubert CC, Eiland EC, Storrow AB, Miller KF, Collins SP. Self-care barriers reported by emergency department patients with acute heart failure: A sociotechnical systems-based approach. Annals of emergency medicine. 2015; 66:1–12. [PubMed: 25616317]

29. Holden RJ, Valdez RS, Schubert CC, Thompson MJ, Hundt AS. Macroergonomic factors in the patient work system: Examining the context of patients with chronic illness. Ergonomics. 2017; 60:26–43. [PubMed: 27164171]

30. Holden RJ, Carayon P, Gurses AP, Hoonakker P, Hundt AS, Ozok AA, Rivera-Rodriguez AJ. SEIPS 2.0: A human factors framework for studying and improving the work of healthcare professionals and patients. Ergonomics. 2013; 56:1669–1686. [PubMed: 24088063]

31. National Research Council, Health Care Comes Home: The Human Factors, National Academies Press. Board on Human-Systems Integration. Division of Behavioral and Social Sciences and Education; Washington, DC: 2011. Committee on the Role of Human Factors in Home Health Care.

32. Sabate, E. Adherence to long-term therapies: Evidence for action. World Health Organization; Geneva, Switzerland: 2003. http://www.who.int/chp/knowledge/publications/adherence_full_report.pdf?ua=1

33. LS R. The CES-D scale: A self-report depression scale for research in the general population. Applied Psychological Measurement. 1997; 1:385–401.

34. Saliba D, Elliot M, Rubenstein L, Solomon D, Young R, Kamberg C, Roth C, MacLean C, Shekelle P, Sloss E, Wenger N. The Vulnerable Elders Survey: A tool for identifying vulnerable older people in the community. J Am Geriatr Soc. 2001; 49:1691–1699. [PubMed: 11844005]

35. Kroenke K, Spitzer R, Williams J, Lowe B. The patient health questionnaire somatic, anxiety, and depressive symptom scales: A systematic review. General Hospital Psychiatry. 2010; 32:345–359. [PubMed: 20633738]

36. Chew L, Bradley K, Boyko E. Brief questions to identify patients with inadequate health literacy. Family Medicine. 2004; 36:588–594. [PubMed: 15343421]

37. Smith M, Wallston K, Smith C. The development and validation of the Perceived Health Competence Scale. Health Education Research. 1995; 10:51–64. [PubMed: 10150421]

38. Ende J, Kazis L, Ash A, Moskowitz MA. Measuring patients' desire for autonomy: Decision making and information-seeking preferences among medical patients. Journal of General Internal Medicine. 1989; 4:23–30. [PubMed: 2644407]

39. Deber R, Kraetschmer N, Irvine J. What role do patients wish to play in treatment decision making? Archives of Internal Medicine. 1996; 156:1414–1420. [PubMed: 8678709]

40. Fagerlin A, Zikmund-Fisher B, Ubel P, Jankovic A, Derry H, Smith D. Measuring numeracy without a math test: Development of the Subjective Numeracy Scale. Medical Decision Making. 2007; 27:672–680. [PubMed: 17641137]

41. Sinclair V, Wallston K. The development and psychometric evaluation of the Brief Resilient Coping Scale. Assessment. 2004; 11:94–101. [PubMed: 14994958]

42. Centers for Disease Control and Prevention (CDC). Behavioral risk factor surveillance system. 2010

43. Hall M, Zheng B, Dugan E, Camacho F, Kidd K, Mishra A, Balkrishnan R. Measuring patients' trust in their primary care providers. Medical Care Research and Review. 2002; 59:293–318. [PubMed: 12205830]

44. Coleman E, Mahoney E, Parry C. Assessing the quality of preparation for posthospital care from the patient's perspective: The care transitions measure. Medical Care. 2005; 43:246–255. [PubMed: 15725981]

45. Marmot MG, Fuhrer R, Ettner SL, Marks NE, Bumpass LL, Ryff CD. Contribution of psychosical factors to socioeconomic differences in health. Milbank Q. 1998; 76:403–448. [PubMed: 9738169]

46. Mitchell P, Powell L, Blumenthal J, Norten J, Ironson G, Pitula C, Rogers R, Froelicher E, Sivarajan R, Czajkowski S, Youngblood M, Huber M, Berkman L. A short social support measure for patients recovering from myocardial infarction: The ENRICHD Social Support Inventory. Journal of Cardiopulmonary Rehabilitation. 2003; 23:398–403. [PubMed: 14646785]

47. Paxton AE, Strycker LA, Toobert DJ, Ammerman AS, Glasgow RE. Starting the conversation: Performance of a brief dietary assessment and intervention tool for health professionals. Am J Prev Med. 2011; 40:67–71. [PubMed: 21146770]

48. Kripalani S, Risser J, Gatti M, Jacobson T. Development and evaluation of the adherence to refills and medications (ARMS) scale among low-literacy patients with chronic disease. Value in Health. 2009; 12:118–123. [PubMed: 19911444]

49. Hall M, Zheng B, Dugan E, Camacho F, Kidd K, Mishra A, Balkrishnan R. Measuring patients' trust in their primary care providers. Med Care Res Rev. 2002; 14:293–318.

50. Shea J, Micco E, Dean L, McMurphy S, Schwartz J, Amstrong K. Development of a Revised Health Care System Distrust Scale. J Gen Intern Med. 2008; 23:727–732. [PubMed: 18369678]

51. Ward JH. Hierarchical Grouping to Optimize an Objective Function. J Am Statistical Assoc. 1963; 58:236–244.

52. von Borries G, Wang H. Partition clustering of high dimensional low sample size data based on p-values. Computational Statistics & Data Analysis. 2009; 53:3987–3998.

53. Liu Y, Hayes DN, Nobel A, Marron J. Statistical significance of clustering for high-dimension, low–sample size data. Journal of the American Statistical Association. 2008; 103:1281–1293.

54. Ding Y, Dang X, Peng H, Wilkins D. Robust clustering in high dimensional data using statistical depths. BMC Bioinformatics. 2007; 8 https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-1188-S1187-S1188.

55. Cali ski T, Harabasz J. A dendrite method for cluster analysis. Communications in Statistics. 1974; 3:1–27.

56. Duda, RO., Hart, PE., Stork, DG. Pattern Classification. 2. John Wiley & Sons; New York: 2001.

57. Tomarken A, Serlin R. Comparison of ANOVA alternatives under variance heterogeneity and specific noncentrality structures. Psychological Bulletin. 1986; 99:90–99.

58. Finkelstein, J., Knight, A., Marinopoulos, S., Gibbons, MC., Berger, Z., Aboumatar, H., Wilson, RF., Lau, BD., Sharma, R., Bass, EB. Evidence Report/Technology Assessment No. 206. AHRQ Publication No. 12-E005-EF., Agency; Rockville, MD: 2012. Enabling patient-centered care through health information technology.

59. U.S. Department of Health and Human Services. Report to Congress: Aging Services Technology Study. Washington, DC; 2012. https://aspe.hhs.gov/basic-report/reportcongress-aging-services-technology-study

60. Dyer, E., Kansagara, D., Mclnnes, D., Freeman, M., Woods, S. Mobile applications and internet-based approaches for supporting non-professional caregivers: a systematic review. VA-ESP Project #05-225., Department of Veterans Affairs; Washington, DC: 2012.

61. Mitzner TL, McBride SE, Barg-Walkow LH, Rogers WA. Self-management of wellness and illness in an aging population. Reviews of Human Factors and Ergonomics. 2013; 8:277–333.

62. Faiola A, Holden RJ. Consumer health informatics: Empowering healthy-lifestyle-seekers through mHealth. Progress in Cardiovascular Diseases. 2017; 59:479–486. [PubMed: 28038910]

63. Vermeulen J, Verwey R, Hochstenback LM, van der Weegen S, Man YP, de Witte LP. Experiences of multidisciplinary development team members during user-centered design of telecare products and services: A qualitative study. Journal of Medical Internet Research. 2014; 16 http://www.jmir.org/2014/2015/e2124/.

64. Abedtash H, Holden RJ. Systematic review of the effectiveness of health-related behavioral interventions using portable activity sensing devices (PASDs). Journal of the American Medical Informatics Association. 2017 doi.org/10.1093/jamia/ocx1006.

65. Nielsen, L. Personas - User Focused Design. Springer; London: 2013.

66. Floyd IR, Cameron Jones M, Twidale MB. Resolving incommensurable debates: a preliminary identification of persona kinds, attributes, and characteristics. Artifact. 2008; 2:12–26.

67. Matthews, T., Judge, T., Whittaker, S. How do designers and user experience professionals actually perceive and use personas?; Proceedings of the SIGCHI conference on human factors in computing systems, ACM; 2012. p. 1219-1228.

68. Blindheim, J., Wulvik, A., Steinert, M. Using Secondary Video Material for User Observation in The Needfinding Process for New Product Development and Design; DS 84: Proceedings of the DESIGN 2016 14th International Design Conference; 2016. p. 1845-1854.

69. Zhang, X., Brown, H-F., Shankar, A. Data-driven Personas: Constructing Archetypal Users with Clickstreams and User Telemetry; Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, ACM; 2016. p. 5350-5359.

70. Phillips JC, Rowsell DJ, Boomer J, Kwon J-Y, Currie LM. Personas to guide understanding traditions of gay men living with HIV who smoke. Qualitative Health Research. 2015; 26:41–54. [PubMed: 25881965]

71. Turner AM, Reeder B, Ramey J. Scenarios, personas and user stories: User-centered evidence-based design representations of communicable disease investigations. Journal of Biomedical Informatics. 2013; 46:575–584. [PubMed: 23618996]

72. Friess, E. Personas and decision making in the design process: an ethnographic case study; Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM; 2012. p. 1209-1218.

73. Burrows A, Gooberman-Hill R, Coyle D. Empirically derived user attributes for the design of home healthcare technologies. Personal and Ubiquitous Computing. 2015; 19:1233–1245.

74. Srinivas P, Cornet V, Holden RJ. Human factors analysis, design, and testing of Engage, a consumer health IT application for geriatric heart failure self-care. International Journal of Human-Computer Interaction. 2017; 33:298–312.

75. Williams I, Brereton M, Donovan J, McDonald K, Millard T, Tam A, Elliott JH. A collaborative rapid persona-building workshop: creating design personas with health researchers. International Journal of Sociotechnology and Knowledge Development (IJSKD). 2014; 6:17–35.

76. Scandurra I, Sjolinder D. Participatory design with seniors: Design of future services and iterative refinements of interactive eHealth services for old citizens. Medicine 2.0. 2013; 2 http://www.medicine20.com/2013/2012/e2012/.

77. Fore D, Goldenhar L, Margolis P, Seid M. Using goal-directed design to create a novel system for improving chronic illness care. JMIR Res Protocols. 2013; 2 http://www.researchprotocols.org/2013/2012/e2043/.

78. Anderson, M. Technology Device Ownership: 2015. Pew Research Center; 2015. Available at: http://www.pewinternet.org/2015/10/29/technology-device-ownership-2015

79. Perrin, A. Social Media Usage: 2005–2015. Pew Research Center; 2015. Available at: http://www.pewinternet.org/2015/10/29/technology-device-ownership-2015

80. Perrin, A., Duggan, M. Americans' Internet Access: 2000–2015. Pew Research Center; 2015. Available at: http://www.pewinternet.org/2015/10/29/technology-device-ownership-2015

81. Vincent CJ, Blandford A. The challenges of delivering validated personas for medical equipment design. Applied Ergonomics. 2014; 45:1097–1105. [PubMed: 24518652]

Highlights

- Data from older adults with heart failure cluster into 6 biopsychosocial personas

- Personas differ on factors such as age, health literacy, and social context.

- Persona descriptions provide an easy reference guide for informatics designers.

Summary

What was already known on the topic

- Personas are a recommended method for user-centered design of information technology, including eHealth products.

- eHealth personas development studies have used qualitative methods to categorize patients along dimensions such as demographics, education, occupation, needs, and attitudes.

- Personas of patients can help eHealth designers create products that accommodate the range or are customized to user differences, e.g., rural vs. and urban dwelling.

What this study added to our knowledge

- Personas of older patients with heart failure differ on a variety of design-relevant dimensions, including age, sex, comorbid diabetes, exhaustion, household work ability, hygiene care ability, physical ability, depression, health literacy, numeracy, Internet availability, contact with home healthcare, trust in providers, informal caregiver support, cohabitation, marital status, and employment status.

- Quantitative personas development methods produce statistically distinct categories of people. These personas can be systematically tested and compared as well as reproduced or modified across eHealth projects and user populations.

- A ten-step personas development method is provided for future use during health informatics design.
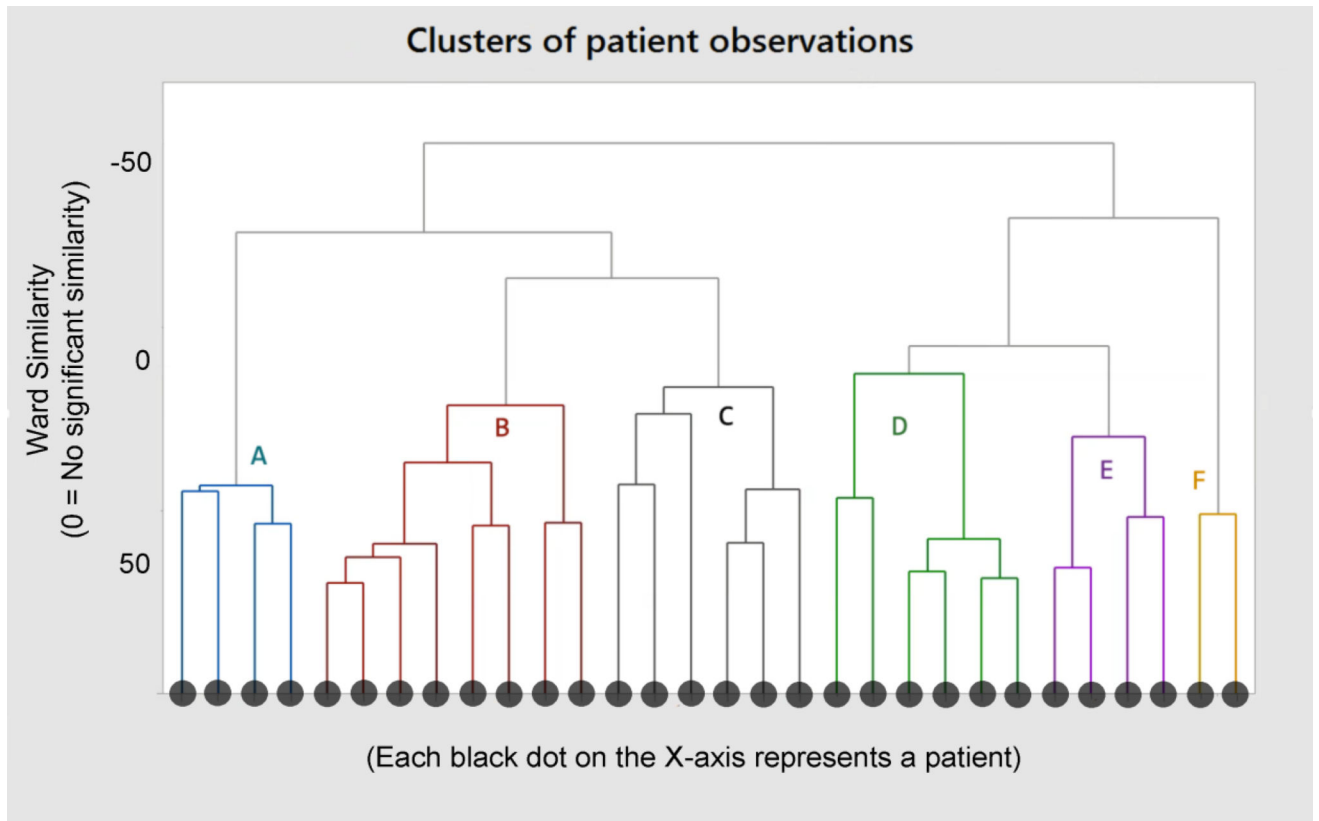
Figure 1.
Dendrogram of six clusters, A through F.

Table 1

Domains and subdomains used in personas development analyses.

Domain	Subdomain
DEMOGRAPHIC	Age [26]
	Education [26]
	Race [26]
	Sex [26]
MEDICAL	Diabetes Comorbidity [26]
	Number of Medications Taken, from electronic medical record
FUNCTIONAL	Exhaustion [33]
	Household Work Ability [34]
	Hygiene Care Ability [34]
	Physical Activity Ability [34]
PSYCHOLOGICAL	Depression Comorbidity [35]
	Subjective Health Literacy [36]
	Health Self-Efficacy [37]
	Healthcare Locus of Control [38] [39]
	Numeracy [40]
	Resilient Coping [41]
TECHNOLOGICAL	Internet Availability [26]
HEALTH BEHAVIOR	Alcohol Usage [42]
HEALTHCARE SYSTEM	Access to Primary Care [26]
	Visited by Home Healthcare Professional [26]
	Trust in Hospital Providers [43]
SOCIAL CONTEXT	Caregiver Support in Transition [44]
	Cohabitation Status [26]
	Marital Status [26]
	Number of Family Members in Contact [45]
	Number of Friends in Contact [45]
	Number of Neighbors in Contact [45]
	Social Support [46]
ECONOMIC CONTEXT	Economic-Related Stress [26]
	Employment Status [26]
OUTCOMES (not used as factors in cluster analysis)	Brain Naturietic Peptide (BNP) value (heart failure biomarker), from electronic medical record
	Dietary Adherence [47]
	Excess (unscheduled) Utilization of Healthcare Services (from electronic medical record)
	Exercise & Rehabilitation [26]
	Medication Adherence [48]
	Mortality, 39 Month, from electronic medical record

Domain	Subdomain
	Use of Patient Portal System [26]

Table 2

Results of cluster comparisons on clustering variables (a–i) and outcome variables (j).

Variable	Test	Test value	*P*-value
a. DEMOGRAPHIC			
Age	WX	5.68	.016
Education	WX	1.52	.296
Race	KW	1.16	.949
Sex	KW	15.19	< .001
b. MEDICAL			
Diabetes Comorbidity	KW	15.95	< .001
Number of Medications Taken	WX	1.93	.195
c. FUNCTIONAL			
Exhaustion	WX	3.05	.089
Household Work Ability	WX	23.25	< .001
Hygiene Care Ability	KW	27.00	< .001
Physical Activity Ability	WX	26.00	< .001
d. PSYCHOLOGICAL			
Depression Comorbidity	WX	4.60	.034
Subjective Health Literacy	WX	7.69	.004
Health Self-Efficacy	WX	2.40	.665
Locus of Control in Healthcare	WX	0.66	.330
Numeracy	WX	2.81	.094
Resilient Coping	WX	1.78	.236
e. TECHNOLOGICAL			
Internet Availability	KW	27.00	< .001
f. HEALTH BEHAVIOR			
Alcohol Usage	WX	.890	.506
g HEALTHCARE SYSTEM			
Access to Care	KW	0.04	> .999
Visited by Home Healthcare Professional	KW	27.00	< .001
Trust in Hospital Providers	WX	2.52	.101
h. SOCIAL CONTEXT			
Caregiver Support in Transition	KW	26.00	< .001
Cohabitation	KW	17.22	.004
Marital Status	KW	14.03	.015
Number of Family Members in Contact	WX	0.57	.725
Number of Friends in Contact	WX	1.05	.439
Number of Neighbors in Contact	WX	0.82	.568
Social Support	WX	2.17	.158
i. ECONOMIC CONTEXT			

Variable	Test	Test value	*P*-value
Economic-Related Stress	WX	2.52	.146
Employment Status	KW	10.56	.061
j. OUTCOMES			
Brain Naturietic Peptide (BNP) Level	WX	0.86	.551
Dietary Adherence	WX	0.98	.498
Excess Utilization of Healthcare Services	WX	1.38	.309
Exercise & Rehabilitation Adherence	WX	8.18	.003
Medication Adherence	WX	.650	.675
Mortality, 39 month	KW	2.64	.755
Patient Portal Usage, Purposeful Use	WX	1.00	.436

KW = Kruskal-Wallis Analysis of Variance (ANOVA); WX = Welch's Chi-Squared (aka Welch's ANOVA).

Table 3

Attributes of the six personas.

Description of factor/outcome	A - "Angela" n=4	B - "Barbara" n=8	C - "Charlene" n=6
Age	70	72	75
Sex	Likely Female	Likely Female	**Female**
Race	**White**	**White**	Likely White
Marital Status	**Married**	Likely Married	**Single**
Employment	Unemployed	Unemployed	Possibly employed
Cohabitation	**Lives with others**		Lives alone
Diabetes	**Not diabetic**	Likely diabetic	**Diabetic**
Depression	More depressed	**Most depressed**	
Caregiver Support In Transition	**Has support**	**Has support**	
Health Literacy	**Least health literate**	**Most health literate**	
Numeracy	**Least numerate**		
Physical Activity Ability	**Most able to do activity**	More able to do activity	Less able to do activity
Household Work Ability	**Most able to do housework**	Able to do housework	Less able to do housework
Hygiene	**Poor hygiene ability**	**Good hygiene ability**	**Good hygiene ability**
Internet Availability	**Internet available**	**Internet available**	Internet likely available
Visited By Home Healthcare Professional	**Visited By Home Healthcare Professional**	**Visited By Home Healthcare Professional**	**Visited By Home Healthcare Professional**
Exhaustion			
Exercise & Rehabilitation	Little exercise	**Least exercise**	
Patient Portal Usage	Higher portal use	**Highest portal use**	Lower portal use
Description of factor/outcome	**D - "David" n=6**	**E - "Earl" n=4**	**F - "Frederick" n=2**
Age	69	71	82
Sex	**Male**	Likely Male	**Male**
Race	Likely White	**White**	**White or non-white (labeled as non-white to reflect population)**
Marital Status	Likely Married	**Married**	**Married**
Employment Status	Unemployed	Possibly employed	Unemployed
Cohabitation	**Lives with others**	**Lives with others**	**Lives with others**
Diabetes Comorbidity	**Diabetic**	**Diabetic**	**Diabetic**
Depression Comorbidity	**Least depressed**	Less depressed	Less depressed
Caregiver Support In Transition	**Has support**	**Has support**	**No support**
Health Literacy		More health literate	
Numeracy		More numerate	**Most numerate**
Physical Activity Ability	Less able to do activity	**Least able to do activity**	More able to do activity
Household Work Ability	Less able to do housework	**Worst ability to do housework**	
Hygiene Ability	**Good hygiene ability**	**Good hygiene ability**	**Good hygiene ability**

Internet Availability	**Internet available**	**Internet available**	**No Internet at home**
Visited By Home Healthcare Professional	**Visited By Home Healthcare Professional**	**Visited By Home Healthcare Professional**	**Not visited By Home Healthcare Professional**
Exhaustion	**Least exhausted**		**Most exhausted**
Exercise & Rehabilitation			**Exercises the most**
Patient Portal Usage	Higher portal use	**No portal use**	**No portal use**

An empty cell denotes the variable was not salient for the cluster, with salience defined for continuous variables as an absolute Z score > 0.4 (see Appendix B). Salience is defined for binomial variables as the proximity to either 0 or 1. Bold denotes the highest salience of the attribute for that cluster.

Table 4

Detailed examples of two personas, "Barbara" and "Frederick."

BARBARA	FREDERICK
Age 72, White female, retired	Age 82, Black male, retired
"I'm good at computers for someone my age… in fact, I already use the patient portal to talk to my doctors!"	*"I'm not going to use this portal… but [it could be] … useful… if my son used it for me."*
Medical	**Medical**
A type 2 diabetes mellitus patient, Barbara also takes a higher number of medications than average.	Frederick has type 2 diabetes mellitus.
Functional	**Functional**
Although able to manage her hygiene, Barbara notes she has issues with physical exhaustion.	Frederick does not have significant issues with caring for himself or with his hygiene. He reports being quite exhausted.
Psychological	**Psychological**
Barbara has been depressed lately. She perceives moderate physical health. She has high health literacy compared to other patients.	Frederick considers himself quite knowledgeable about personal health. He rarely feels depressed or down in the dumps.
Technological	**Technological**
Barbara has a computer with Internet access. She already uses the health system's patient portal to communicate with her doctors and nurses and is receptive to new information technologies.	A former computer operator before retiring, Frederick has no desire to learn new technologies. He has no Internet at home and would not use a patient portal unless his son would use it for him.
Social	**Social**
Barbara is married and lives with her husband. Her husband helped care for her after she was discharged from the hospital.	Frederick is married and lives with his wife. He has a high level of contact with friends, and has his son come over to help when he or his wife needs it.
Healthcare System	**Healthcare System**
Barbara has a home healthcare professional visit her regularly.	Other than his physical therapist, Frederick has not been seen by a home healthcare professional.
Health Behaviors	**Health Behaviors**
Exercise is not a daily routine for Barbara.	Frederick exercises quite a bit for physical therapy and rehabilitation. He finds it exhausting.

Table 5

The authors' proposed ten-step process for personas development and use by designers in the domain of health and healthcare.

1.	**Recruit patients (or informal caregivers).** Attempt to maximize variability and sample size (ideally 5 participants per clustering themselves or can provide data on patients. Alternatively, an existing data set could be used.
2.	**Administer standardized surveys.** Structured, quantitative data are needed for cluster-analytic personas generation. Several validated instruments are available, including the ones used in this study. An attempt should be made to cover multiple biopsychosocial domains as efficiently as possible.
3.	**Collect medical information.** If possible, medical record or self-reported medical data can help characterize personas and compare the sample to a broader population.
4.	**Collect additional qualitative data.** If possible, individual interviews or observations can be used to collect rich data for each patient beyond their survey results. These provide context and interpretation of survey results as well as additional information.
5.	**Clean and prepare the data for analysis.** This will include checking raw data, constructing scales, handling missing data, and performing transformations.
6.	**Perform multivariate cluster analysis.** A variety of approaches clustering, k-means partitioning, expectation maximization clustering) and software packages (e.g., MINITAB, SAS, R) can be used.
7.	**Compare clusters on variables.** The variables can include both those used in the clustering as well as others hypothesized to differ between clusters but not included in cluster analysis. This step is used to test cluster validity and identify important differences between clusters. For small sample sizes, non-parametric analyses are recommended.
8.	**Create tables, figures, and narratives depicting personas.** Use results to create depictions of personas. These should be quick and easy to understand, to avoid overwhelming designers.
9.	**Add context and "life" to the through persona documents.** This could be as simple as names and photos, but could also include quotes or observations from qualitative data collection or the relevant literature. Personas can be given a rich profile including history, family, and other details. More inventive personas may include a video, a website, a deck of cards, or an acted-out skit. Designers should be able to print or otherwise easily access persona documents during design to imagine the kind of user for whom they are designing.
10.	**Perform a validation of the personas.** Although resource- and time-intensive, a follow-up can test whether empirically-derived personas can be distinguished in a new sample. If different designs are created to better accommodate specific personas, the design-persona combinations can be tested (e.g., do people matching Persona A prefer Design A, whereas those matching Persona B prefer Design B?)