

Persona Development in WA State: A Mixed Methods Approach Using Statewide Survey Data

Jordan M Garcia, Rebecca A Hills, Debra Revere, Jaewon Lim, Adam S Elder,
Chris Baumgartner, Bryant T Karras, Janet G Baseman

Submitted to: Online Journal of Public Health Informatics
on: April 21, 2025

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5

Supplementary Files..... 26

 Figures 27

 Figure 1..... 28

 Figure 2..... 29

 Figure 3..... 30

 Multimedia Appendixes 31

 Multimedia Appendix 1..... 32

 Multimedia Appendix 2..... 32

 Multimedia Appendix 3..... 32

Persona Development in WA State: A Mixed Methods Approach Using Statewide Survey Data

Jordan M Garcia¹; Rebecca A Hills² PhD; Debra Revere³ PhD; Jaewon Lim⁴; Adam S Elder² PhD; Chris Baumgartner⁵; Bryant T Karras⁵ PhD; Janet G Baseman² PhD

¹ School of Social Work University of Washington Seattle US

² Department of Epidemiology School of Public Health University of Washington Seattle US

³ Department of Health Systems & Population Health School of Public Health University of Washington Seattle US

⁴ Department of Biostatistics School of Public Health University of Washington Seattle US

⁵ Washington State Department of Health Tumwater US

Corresponding Author:

Jordan M Garcia

School of Social Work
University of Washington
4101 15th Ave NE
Seattle
US

Abstract

Background: Personas, fictional profiles representing user segments, play an important role in human-centered design, ensuring tools are tailored to the needs of users. Although public health organizations often develop information systems to promote population health, human-centered design methods and personas are generally underutilized in public health informatics projects.

Objective: This study presents a novel, mixed-methods approach to developing data-driven personas for use in public health information system design, leveraging two statewide surveys conducted in Washington (WA) State. The aim is to produce realistic, representative, and actionable personas that reflect the diversity of a state population and support user-centered design in public health initiatives.

Methods: Quantitative (cluster analysis) and qualitative (thematic review and quote extraction) methods were applied to two statewide survey datasets: 1) a statewide Knowledge, Attitudes, and Practices (KAP) survey (N=1,103) which employed random, address-based sampling, and 2) a subset of the KAP respondents (N=143) which included more targeted questions on opinions and preferences related to public health information systems. Characteristics examined included demographics, technological readiness, opinions about public health policies, and experience using online health tools.

Results: K-prototype clustering resulted in five clusters. These five clusters were studied using both quantitative and qualitative analysis of key factors of the WA State population to build 13 personas. Each persona represents a different population demographic, varying levels of technological readiness and attitudes toward public health policies, and differing experiences with online health tools. Persona descriptions are further elucidated with a short profile and 2-3 quotes.

Conclusions: This study offers a scalable and adaptable framework for persona development in public health, demonstrating how existing datasets can be transformed into effective design tools. Through a mixed-methods approach, personas that reflect the diverse needs, preferences, and behaviors of WA State residents were created. These personas can enhance the design, development, and evaluation of public health information systems by centering user experience. Persona development and the methods described here can be used in future public health informatics projects to assist in formative research, guide design and development, inform usability testing, and shape communication strategies. By bridging the gap between large-scale data and user-centered design, this approach provides a practical model for making public health technologies more aligned with community needs.

(JMIR Preprints 21/04/2025:75422)

DOI: <https://doi.org/10.2196/preprints.75422>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in [http://www.jmir.org](#)

No. Please do not make my accepted manuscript PDF available to anyone. I understand that if I later pay to participate in [https://www.jmir.org](#)

Original Manuscript

Jordan M Garcia,¹ Rebecca A Hills,² Debra Revere,³ Jaewon Lim,⁴ Adam S Elder,² Chris Baumgartner,⁵ Bryant T Karras,⁵ Janet G Baseman²

¹ School of Social Work, University of Washington, Seattle, WA, United States

² Department of Epidemiology, School of Public Health, University of Washington, Seattle, WA, United States

³ Department of Health Systems & Population Health, School of Public Health, University of Washington, Seattle, WA, United States

⁴ Department of Biostatistics, School of Public Health, University of Washington, Seattle, WA, United States

⁵ Washington State Department of Health, Tumwater, WA, United States

Persona Development in WA State: A Mixed Methods Approach Using Statewide Survey Data

Abstract

Background:

Personas, fictional profiles representing user segments, play an important role in human-centered design, ensuring tools are tailored to the needs of users. Although public health organizations often develop information systems to promote population health, human-centered design methods and personas are generally underutilized in public health informatics projects.

Objective:

This study presents a novel, mixed-methods approach to developing data-driven personas for use in public health information system design, leveraging two statewide surveys conducted in Washington (WA) State. The aim is to produce realistic, representative, and actionable personas that reflect the diversity of a state population and support user-centered design in public health initiatives.

Methods:

Quantitative (cluster analysis) and qualitative (thematic review and quote extraction) methods were applied to two statewide survey datasets: 1) a statewide Knowledge, Attitudes, and Practices (KAP) survey (N=1,103) which employed random, address-based sampling, and 2) a subset of the KAP respondents (N=143) which included more targeted questions on opinions and preferences related to public health information systems. Characteristics examined included demographics, technological readiness, opinions about public health policies, and experience using online health tools.

Results: K-prototype clustering resulted in five clusters. These five clusters were studied using both quantitative and qualitative analysis of key factors of the WA State population to build 13 personas. Each persona represents a different population demographic, varying levels of technological readiness and attitudes toward public health policies, and differing experiences with online health tools. Persona descriptions are further elucidated with a short profile and 2-3 quotes.

Conclusions:

This study offers a scalable and adaptable framework for persona development in public health, demonstrating how existing datasets can be transformed into effective design tools. Through a mixed-methods approach, personas that reflect the diverse needs, preferences, and behaviors of WA State residents were created. These personas can enhance the design, development, and evaluation of public health information systems by centering user experience. Persona development and the

methods described here can be used in future public health informatics projects to assist in formative research, guide design and development, inform usability testing, and shape communication strategies. By bridging the gap between large-scale data and user-centered design, this approach provides a practical model for making public health technologies more aligned with community needs.

Keywords: Personas; user-centered design; public health informatics; information system design; cluster analysis; qualitative data

Introduction

Public health informatics projects face the challenge of designing tools that resonate with diverse populations while being applied universally [1] [2]. Due to time pressures and limited resources of public health work, it is often difficult to gather sufficient user input to gain a nuanced understanding of the needs of the population being served. Further, surveys and population summaries may not adequately capture the complexity and variability in public attitudes and behaviors [3] [4]. User centered approaches can help bridge this gap by focusing on the specific needs of individuals [5]. One such user centered design (UCD) tool, personas, can assist in representing individual needs [6]. This study focuses on the creation of detailed personas using a rich dataset, offering a novel approach to public health informatics design work that balances empirical data with granular insights.

Personas are fictional representations of individual users and may be best known for their use in marketing; however, personas are also used by informaticians, designers, and developers to better understand the needs, behaviors, and preferences of potential users of their products, systems, or services [7]. Generally, personas may include demographic information and behavioral patterns derived from user research. As part of a UCD approach, personas foster empathy with users, generating valuable insights for product design, development, and strategy [6]. Essentially, personas can provide a more concrete and actionable representation of both unstructured and complex user data, making it easier to interpret.

Although personas are widely used in information systems design, their application in health information technology (IT) and public health informatics remains limited [8] [9] [10] [11]. Public health efforts, in particular, could benefit from the use of personas, especially in contexts where resources to support direct user engagement and data collection are constrained. In public health, personas may be useful in various processes: the design and evaluation of products, recruiting for usability testing activities [12], and facilitating communication across stakeholder groups during development and decision-making processes [13]. In particular, personas can be an innovative strategy for public health teams by offering relatable, data-driven representations of target populations. This approach can enhance the reach and effectiveness of public health initiatives by keeping user needs at the forefront of design and implementation. Additionally, personas can support stronger communication and marketing efforts, making public health campaigns more engaging and accessible to diverse audiences. By integrating personas, teams may create more responsive, user-centered solutions that can improve public health outcomes.

Traditional methods of persona development often rely on small sample sizes or limited information, leading to less accurate and less representative user profiles [14]. Additionally, there is no standardized methodology, resulting in inconsistencies and differing priorities in the development process [15]. Some approaches prioritize rich qualitative data, obtaining the data from a very limited sample of the population. While these processes are designed to capture detailed individual

experiences and perspectives, they may miss insights from unsampled portions of the population or key overarching trends [16]. In contrast, development processes focused on qualitative methods may use larger pools of user data, but generalizations may occur, and subtleties can be lost [17] [18]. Mixed methods approaches seek to find a balance between these two [6]. This study emphasizes the importance of a comprehensive approach that integrates both robust quantitative metrics and detailed qualitative insights to create a more extensive and nuanced understanding of user experiences and needs.

The objective of this study is to develop a set of personas for use in public health informatics initiatives using a mixed-methods approach. By synthesizing quantitative and qualitative data from representative datasets, this study aims to create realistic and representative personas that can inform the design and implementation of public health technologies. These personas are intended to support various project stages, including formative research, usability testing, and communication strategies, ensuring that public health technologies are tailored to meet the specific needs, preferences, and barriers faced by diverse population segments. This study demonstrates the efficacy of integrating quantitative and qualitative data in persona development, leveraging an approach that combines statistical rigor with contextual insights. The resulting approach provides a versatile template adaptable for other public health initiatives.

Materials and methods

Context

Washington State's Verifiable Clinical Information system, WA Verify, is a free, smartphone-based tool that was launched in November 2021 to provide convenient access to COVID-19 vaccine records. To evaluate WA Verify, a statewide survey was conducted to assess WA residents' knowledge, attitudes, and practices regarding public health technologies, with a specific focus on WA Verify. In general, the survey found high rates of smartphone and internet access as well as concerns about data privacy and the reliability of digital health records.

As pandemic restrictions eased, a second-phase survey was conducted to explore potential expanded uses for WA Verify and its underlying technology. This survey examined preferences and concerns related to privacy and security. Findings revealed that participants generally trusted public health tools but worried about the security of their personal health data. Nevertheless, many participants acknowledged the societal benefits of these tools, ultimately weighing the advantages over perceived risks.

The two rich datasets described above offered a unique opportunity to understand the WA State population as a whole and capture some of the specific concerns and opinions of WA residents in their own words. In public health, time and resource limitations may hinder optimal user input during the planning and design of informatics projects. To address this gap, personas emerged as one way to leverage the available data in order to represent users and succinctly communicate their needs and preferences. A persona is a fictitious individual created to represent a user group, providing designers with personal stories that offer insights into motivations and behaviors. While traditional UCD efforts often rely on small sample sizes and qualitative data, the two evaluation survey datasets with rich representative data as well as focused qualitative data on informatics topics, presented an opportunity to create a set of data-derived personas. The goal was to design personas that exhibit the following attributes:

1. **Simple / Relatable:** Provide a collection of distinct individuals who can be understood at an

intuitive and personal level to help in product design.

2. **Accurate:** As much as possible, accurately represent the relationship between demographic characteristics and technology-related tendencies.
3. **Diverse:** Represent a range of characteristics. While not capturing every individual, the set of personas should encapsulate diverse attributes, encouraging consideration of various population segments.
4. **Actionable:** Provide insights that can inform decisions and strategies in public health informatics projects.
5. **Comprehensive:** Elucidate motivations, intentions, and needs, not just demographics.

In the development process, these persona attributes may be at odds with one other. For example, the goals of relatability and actionability may conflict with complete statistical accuracy. This approach seeks to find a balance among these goals in order to create a data-driven and useful set of personas.

Data Collection

The WA Verify evaluation project conducted one statewide survey and later a second-phase survey; data from these two surveys were used for persona development. Survey methods are summarized below and are described in detail elsewhere [19].

The statewide survey aimed to gather insights on technology and public health attitudes from a broad sample of WA residents. Survey invitations were mailed to a simple random sample of 5,000 WA households. The 32 survey questions covered technology use, digital literacy, internet access, COVID-19 verification experiences, attitudes toward digital health tools, and opinions on public health policies. The complete statewide survey instrument is included in Appendix 1. Data collection took place between September 2022 and January 2023. Of the 5,000 invitations mailed, 1,491 households responded while 302 invitations were returned as undeliverable, resulting in a 32% response rate.

The second-phase survey targeted a smaller sample but aimed to gather more specific information about barriers to technology use in the public health context. The survey had 53 questions and focused on digital literacy, internet safety perceptions, digital privacy and security, confidence in public health initiatives, and opinions on hypothetical public health tools including some space for free responses from participants to elaborate on their views. The complete second-phase survey instrument is included in Appendix 2. The survey sample consisted of 321 respondents from the statewide survey who expressed interest in future studies, approximately 20% of all initial respondents. Data collection took place in April 2023, yielding responses from 143 individuals and resulting in a response rate of 44.5%.

Data Analysis

Personas were developed from the two survey datasets using a mixed-methods approach, combining quantitative and qualitative analyses. While cluster analysis provided a statistical foundation, ensuring that the final product was aligned with the survey data, this method alone did not allow for comprehensive personas. The mixed methods approach allowed for refinement of the initial clusters into distinct groups with specific demographic characteristics, aiming to capture specific needs of diverse user groups. The process involved five main steps: cluster analysis, cluster segmentation, persona profile development, persona card creation, and validation. Error: Reference source not found provides a graphical representation of this process, and each step is summarized below.

1. **Cluster Analysis:** Cluster analysis identified five primary clusters based on 14 key variables including demographic, technological, and policy opinion characteristics from the WA Statewide Survey dataset. An additional cluster was created to represent nonbinary and transgender individuals who did not have sufficient representation to be included in the primary cluster analysis.
2. **Cluster Segmentation:** Quantitative methods expanded the five primary clusters into 14 subclusters or cluster segments. Joint distributions of age, gender, and race/ethnicity were examined to create distinct cluster segments within each cluster. Outcome variables were then summarized and assigned to these segments, creating personas.
3. **Persona Profile Development:** Qualitative data integration enriched personas with quotes and themes extracted from both survey datasets. This process added depth and nuance to personas with additional characteristics and narrative descriptions. Redundant personas were consolidated, resulting in a final set of 13 personas.
4. **Persona Card Creation:** After defining the personas, persona cards were developed to visually represent each persona's characteristics, needs, and motivations.
5. **Validation:** A validation process ensured that the personas accurately reflected the demographic and attitudinal diversity observed in the WA datasets. Team review and quantitative comparison methods were used to iteratively refine and validate each persona.

[Figure 1. Steps of the persona development process.]

Quantitative Methods

Cluster Analysis

To identify distinct groups within the dataset, K-prototype clustering with Gower's distances [20] was used. K-prototype clustering is an extension of K-means that is suitable for datasets with numerical and categorical variables [21] [22]. A distance function was defined to create clusters based on variables of interest, producing a "distance" value indicating similarity between respondents, with smaller values denoting greater similarity and larger values less similarity. The 14 variables used in this analysis can be seen in Table 1.

Table 1. Cluster analysis variables

Variable (Abbreviation)	Description	Cluster Analysis Categories
Demographic		
Age	What is your age?	18-39; 40-59; 60+
Gender	Are you _____?	Female; Male; Transgender; Non-binary/non-conforming; Prefer not to respond
Race, Ethnicity	How would you describe your race and ethnicity?	White; Black; Hispanic; Asian; Another race or ethnicity
Geographic Region (Region)	FIPS Code	Eastern Washington; Western Washington
English Only at Home (English Only)	Do you speak a language other than English at home?	Only English; Language Other Than English
Parental Status (Parent)	Are you a parent or guardian to a child under 18 years old?	Yes parent/guardian to <18; No parent/guardian to <18
Education	What is your highest level of education?	2-year degree/some college; 4-year degree or more; HS graduate; No HS degree
Technological / Non-demographic Variables		
Technological Readiness (Tech Readiness)	Calculated Variable: Overall, how confident, if at all, do you feel using computers,	Higher tech readiness; Lower tech readiness

	smartphones, or other electronic devices to do the things you need to do online?; How true is the following statement? "When I get a new electronic device, or need to do a new task on it, I usually need someone else to set it up, show me how to use it, or help me."	
eHealth Tool Usage	Have you ever used any electronic health tools?	eHealth tool usage; No eHealth tool usage
Method of Learning of WA Verify (WA Verify Refer)	How did you first hear about WA Verify?	Healthcare; Never Heard; News; Other; Referral
WA Verify Usage	Calculated Variable: Do you use WA Verify?; How willing would you be to use a portable electronic COVID-19 vaccine record?	Yes; Willing; Not Willing
COVID Policy Opinion (Opinion Covid Pol.)	How do you feel about policies that require proof of vaccination or a negative COVID-19 test result to enter spaces that are high risk of COVID-19 spread?	Oppose; Support
Health Portal Usage (Health Portal)	Have you ever used any electronic health tools?: An online patient medical record or health portal (for example MyChart).	Yes; No
WA Notify Usage	Have you ever used any electronic health tools?: WA Notify (WA's COVID-19 Exposure Notification Tool).	Yes; No
Tracking Apps Usage (Tracking)	Have you ever used any electronic health tools?: A tool for keeping track of my activities like diet and exercise (for example Fitbit, Strava, or MyFitnessPal)	Yes; No

Weighting was applied as a part of the cluster analysis to emphasize the importance of technology-related variables. On the first round of analysis, equal weights were applied to each variable; however, this approach resulted in most clusters having high technology proficiency (e.g., high WA Verify usage and technology readiness). While this suggests a generally technology savvy population, it does not provide useful insights for targeting subgroups with lower technology readiness. Since the personas were specifically developed in the context of informing design and implementation of digital public health tools, it was crucial to differentiate between groups based on technology-related factors. To address this, technology-related variables were assigned a weight of 5, and demographic variables a weight of 1. This weighting scheme allowed for the creation of more distinct clusters displaying varying levels of technology adoption and readiness. This technique helped to highlight various user groups with different levels of technological readiness.

Construction of the clusters was carried out by maximizing the 50-fold cross-validated silhouette score [23]. The silhouette score, ranging from -1 to 1, measures how similar each data point is to its own cluster compared to other clusters, with higher scores indicating better-defined clusters. This approach enhances the reliability and accuracy of the results by allowing for the assessment of clustering performance on data not used in its construction, reducing overfitting risk, improving cluster stability, and enhancing generalizability across different data subsets. The optimal number of clusters was determined to be five, balancing granularity and meaningful segmentation within the dataset.

Addition of a Supplemental Cluster

Special methods were employed to ensure the inclusion of demographic groups with limited presence in the sample, particularly nonbinary and transgender individuals. Due to the small number of survey respondents identifying as nonbinary or transgender, the primary cluster analysis focused exclusively on respondents identifying as male or female. To address this limitation and ensure representation of the gender diversity present in WA State, an additional, supplemental cluster

(Supplemental Gender Cluster) was created, including only respondents who identified as nonbinary or transgender. The cluster's summary characteristics were defined using an aggregate of the characteristics from the eight respondents who comprised this group.

Cluster Segmentation

Cluster analysis was used to initially identify groups with shared characteristics but fell short of producing usable personas due to the diversity within each cluster. To create personas that accurately reflected the WA State population, it was necessary to further segment the initial clusters. The original five clusters were divided into 13 subgroups based on race/ethnicity, age, and gender, all factors that are understood to influence behaviors and attitudes towards health tools [24] [25] [26] [27] [28] [29]. Segmentation was achieved by comparing each cluster's demographic distribution with the demographics of the overall sample, helping to identify any overrepresented characteristics in the clusters. Overrepresentation suggested the creation of additional subgroups (between one and five per cluster) which became 13 subgroups, or personas.

An example of the segmentation process can be seen in Table 2 which compares the distribution of the entire sample to that of Cluster 1. Cluster 1 had a higher proportion of white, male, 40–59-year-olds as well as Asian, female, 40-59-year olds prompting division into two subclusters, each representing one of these identified groups (white, male, 40-59 and Asian, female, 40-59).

Table 2. Three-way table age of, gender, and race ethnicity among total survey sample (n=1103) and Cluster 1 (n=272) with Observed/Expected for cluster categories

	All		White		Black		Hispanic		Asian		Other	
	Survey	Cluster 1 (O/E)	Survey	Cluster 1 (O/E)	Survey	Cluster1 (O/E)	Survey	Cluster1 (O/E)	Survey	Cluster1 (O/E)	Survey	Cluster1 (O/E)
18-39												
Female	182	46 (102%)	106	27 (103%)	8	2 (101%)	25	5 (81%)	18	6 (97%)	25	6 (97%)
Male	110	22 (81%)	67	12 (73%)	2	0	6	2 (135%)	19	5 (107%)	16	3 (76%)
40-59												
Female	208	74 (144%)	155	53 (139%)	5	0	16	4 (101%)	19	11 (235%)	13	6 (187%)
Male	136	59 (176%)	106	49 (187%)	5	1 (81%)	2	1 (203%)	11	4 (147%)	12	4 (135%)
60+												
Female	259	43 (67%)	225	38 (68%)	1	0	7	0	4	1 (101%)	22	4 (74%)
Male	208	28 (55%)	173	22 (52%)	7	0	1	0	11	3 (111%)	16	3 (76%)

This segmentation process, applied to each of the five quantitatively derived clusters, resulted in 13 distinct personas. When combined with the Supplemental Gender Cluster, which remained a standalone persona without further subdivision, a total of 14 personas were created. Outcome variables of interest (e.g., Technological Readiness, WA Verify Usage, etc.) were identified for each persona. Most subclusters were defined by the demographic characteristics determined during cluster segmentation and the most frequently observed values of the outcome variables of interest within that subcluster.

Additionally, some personas were modified to better capture the diversity of demographics and outcome variables in the population. The distribution of variables in subclusters was compared to the overall statewide population to ensure representativeness. Since each persona can only reflect one value per variable, it is challenging to establish actionable, diverse, and representative personas. To address this, less common subcluster characteristics were sometimes selected over more common characteristics to improve diversity and representation. Further modifications are detailed in the Validation section, ensuring that the personas are data-driven and representative of Washington State's population.

Qualitative Methods

During the qualitative phase of persona development, participant quotes and responses were incorporated to enrich the personas by capturing nuanced perspectives, needs, and experiences within each defined demographic segment. Thematic analysis identified recurring themes across responses, revealing common sentiments. This process added qualitative depth and real-world context to each persona.

Persona Profile Development

Following the completion of cluster analysis and qualitative data review, a synthesis and refinement process was undertaken to further enhance the personas and ensure that they portrayed realistic representations of individuals. Key aspects crucial for constructing comprehensive personas [30] were identified and aligned with best practices for persona development [31] [32], then tailored to incorporate information specifically relevant to public health information system efforts. These aspects included pertinent demographic information, public health opinions, quotes, and drivers (including needs and motivations).

Persona development and refinement took place as an iterative process with three UW team members reviewing personas, examining quantitative and qualitative data, and reviewing and revisiting personas. The following additions and adjustments were made to personas during the iterative refinement process (not necessarily in this order):

- Additional stories were built around the personas based on qualitative survey data from individuals with matching or similar characteristics to the persona;
- Personas were given exact ages, marital statuses, and occupations;
- Quotes and variations on quotes were assigned to personas from survey respondents with similar characteristics;
- New quotes were created where only partial ideas existed in the qualitative data;
- Personas were given names and pictures;
- One persona was removed (cluster 2, persona 7) as seen in Table 4 because it was nearly identical to several other personas and the group determined it was not a useful addition to the set.

During the refinement process, representative quotes and key motivations (drivers) were added to each persona. Quotes were selected during the qualitative analysis and matched to demographically similar personas, with minor adjustments for clarity and brevity. Drivers emerged from thematic analysis and were grounded in survey data. Throughout this process, the balance between sufficient detail and making realistic personas while avoiding stereotypes and oversimplification was carefully considered.

Persona Card Creation

Following the development of the demographic profiles, persona cards were created to serve as visual and textual representations of the personas. The information about each persona derived from the steps previously outlined were translated to a concise format to facilitate their use in design and decision-making processes. First, a standard template was identified to outline the sections for demographics, public health opinions, quotes, and drivers. The persona demographic profiles were transferred into this template. Each persona card also included a representative image which was selected to be demographically appropriate and aligned with the persona's profile. After the card drafts were established, they were reviewed by UW team members, looking for redundancies and

inconsistencies. This process resulted in 13 persona cards.

Validation

Because the process by which the personas were developed was novel, and in some cases ad-hoc, a validation step was used to confirm that the final persona characteristics were reflective of the observed data. Each persona was validated using the following steps:

1. A group of respondents was created for each persona where each record matched the following characteristics exactly: age group, gender, race/ethnicity, geographic region, language at home, parental status, and level of education.
2. In some cases, when this group of respondents was small or non-existent, a second, larger subset was also considered that included both exact matches and individuals that matched on all but one of the seven characteristics mentioned in Step 1.
3. The distributions of technological readiness, opinions on COVID-19 policies, use of digital health portals, and use of health tracking software were tabulated and compared to the characteristics assigned to the given persona.
4. In cases where a mismatch between the persona characteristic and the respondent summary were observed, the personas team discussed and came to a collaborative decision on whether to make a change to the persona characteristic.

Four major discrepancies were identified through this process, and each prompted a discussion among the team. In two cases, the group determined the persona's characteristics should reflect the demographic groupings and changes were made to reflect this. However, in two cases the discrepancy was retained. For both affected personas, a thoughtful decision was made to retain groups who are generally less represented, likely have unique needs, and are not well understood with respect to habits and preferences regarding public health information systems (i.e., low technological readiness and less experience with online health tools). These decisions reflect the belief that although some characteristics are less common in the generally tech-savvy population of WA State, it is important to take less-common needs and traits into consideration to design accessible tools for all WA residents.

Results

Thirteen personas emerged through the development process, twelve derived from the quantitative clustering and expansion process and one Supplemental Gender Cluster.

Clusters

Cluster analysis identified five clusters. Additionally, a supplemental cluster was defined to identify a segment of the survey population not considered in the analysis. Table 3 displays these six clusters.

Table 3. Five quantitatively identified clusters and one supplemental gender cluster with summary characteristics

	All (n = 1103)	Cluster (n = 272)	1 Cluster (n = 223)	2 Cluster (n = 302)	3 Cluster (n = 57)	4 Cluster (n = 249)	5 Supplemental (n = 8)
Age							
18-39	292 (26%)	68 (25%)	97 (43%)	7 (2.3%)	0 (0%)	120 (48%)	5 (63%)
40-59	344 (31%)	133 (49%)	68 (30%)	36 (12%)	17 (30%)	90 (36%)	1 (13%)
60+	467 (42%)	71 (26%)	58 (26%)	259 (86%)	40 (70%)	39 (16%)	2 (25%)
Gender							
Female	649 (59%)	163 (60%)	128 (57%)	164 (54%)	44 (77%)	150 (60%)	0 (0%)
Male	454 (41%)	109 (40%)	95 (43%)	138 (46%)	13 (23%)	99 (40%)	0 (0%)
Nonbinary	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	5 (63%)
Transgender	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (38%)

Race/Ethnicity							
White	832 (75%)	201 (74%)	156 (70%)	251 (83%)	48 (84%)	176 (71%)	5 (63%)
Black	28 (2.5%)	3 (1.1%)	8 (3.6%)	5 (1.7%)	1 (1.8%)	11 (4.4%)	0 (0%)
Hispanic	57 (5.2%)	12 (4.4%)	18 (8.1%)	8 (2.6%)	0 (0%)	19 (7.6%)	1 (13%)
Asian	82 (7.4%)	30 (11%)	21 (9.4%)	13 (4.3%)	4 (7.0%)	14 (5.6%)	1 (13%)
Others	104 (9.4%)	26 (9.6%)	20 (9.0%)	25 (8.3%)	4 (7.0%)	29 (12%)	1 (13%)
Region							
Eastern	205 (19%)	37 (14%)	35 (16%)	69 (23%)	9 (16%)	55 (22%)	1 (13%)
Western	898 (81%)	235 (86%)	188 (84%)	233 (77%)	48 (84%)	194 (78%)	7 (88%)
English							
Only English	969 (88%)	230 (85%)	191 (86%)	281 (93%)	53 (93%)	214 (86%)	8 (100%)
Other Language	134 (12%)	42 (15%)	32 (14%)	21 (7.0%)	4 (7.0%)	35 (14%)	0 (0%)
Parent to Child <18							
Yes	271 (25%)	83 (31%)	60 (27%)	23 (7.6%)	7 (12%)	98 (39%)	0 (0%)
No	832 (75%)	189 (69%)	163 (73%)	279 (92%)	50 (88%)	151 (61%)	8 (100%)
Education							
4-year degree or more	657 (60%)	207 (76%)	146 (65%)	123 (41%)	38 (67%)	143 (57%)	6 (75%)
2-year or some college	290 (26%)	53 (19%)	58 (26%)	101 (33%)	15 (26%)	63 (25%)	1 (13%)
High school graduate	123 (11%)	11 (4.0%)	18 (8.1%)	58 (19%)	4 (7.0%)	32 (13%)	1 (13%)
Less than high school	33 (3.0%)	1 (0.4%)	1 (0.4%)	20 (6.6%)	0 (0%)	11 (4.4%)	0 (0%)
Technological Readiness							
Higher	826 (75%)	256 (94%)	219 (98%)	97 (32%)	13 (23%)	241 (97%)	8 (100%)
Lower	277 (25%)	15 (5.9%)	4 (1.8%)	205 (68%)	44 (77%)	8 (3.2%)	0 (0%)
Method of Learning of WA Verify							
Referral	223 (20%)	136 (50%)	26 (12%)	17 (5.6%)	10 (18%)	34 (14%)	1 (13%)
News	144 (13%)	53 (19%)	11 (4.9%)	31 (10%)	27 (47%)	22 (8.8%)	0 (0%)
Healthcare	117 (11%)	62 (23%)	19 (8.5%)	19 (6.3%)	3 (5.3%)	14 (5.6%)	0 (0%)
Others	47 (4.3%)	17 (6.3%)	11 (4.9%)	8 (2.6%)	3 (5.3%)	8 (3.2%)	1 (13%)
Never heard	572 (52%)	4 (1.5%)	156 (70%)	227 (75%)	14 (25%)	171 (69%)	6 (75%)
WA Verify Usage							
Yes	276 (25%)	234 (86%)	5 (2.2%)	21 (7.0%)	7 (12%)	9 (3.6%)	1 (13%)
Willing	560 (51%)	17 (6.3%)	193 (87%)	150 (50%)	38 (67%)	162 (65%)	6 (75%)
Not willing	267 (24%)	21 (7.7%)	25 (11%)	131 (43%)	12 (21%)	78 (31%)	1 (13%)
Covid Policies							
Oppose	226 (20%)	26 (9.6%)	27 (12%)	69 (23%)	8 (14%)	96 (39%)	0 (0%)
Support	877 (80%)	246 (90%)	196 (88%)	233 (77%)	49 (86%)	153 (61%)	8 (100%)
Health Portal							
	949 (86%)	262 (96%)	212 (95%)	199 (66%)	57 (100%)	219 (88%)	7 (88%)
WA Notify Usage							
	535 (49%)	235 (86%)	223(100%)	28 (9.3%)	49 (86%)	0 (0%)	2 (29%)
Tracking							
	660 (60%)	209 (77%)	181 (81%)	24 (7.9%)	45 (79%)	201 (81%)	6 (75%)

* **Bold** results were larger than expected based on the full sample.

Persona Segments

From the five calculated clusters and the Supplemental Gender Cluster, 14 personas were identified to represent distinct segments of the population based on race/ethnicity, age, and gender. By identifying overrepresentation of these demographic characteristics, personas were defined within each cluster. This expanded cluster-derived persona list is displayed in Table 4.

Table 4. Primary persona list derived from expansion of cluster analysis groupings

Persona Number	Cluster Number	Age Group	Gender	Race/Ethnicity	Region	English Only	Parent	Education	Tech Readiness	WA Verify Referral	WA Verify Usage	Opinion Covid Pol.	Health Portal	WA Notify Usage	Tracking
1	1	40-59	Male	White	W	Yes	No	4-year	High	Referral	Yes	Support	Yes	Yes	Yes
2	1	40-59	Female	Asian	W	No	Yes	4-year	High	Referral	Yes	Support	Yes	Yes	Yes
3	2	18-39	Female	White	W	Yes	No	4-year	High	Referral	Willing	Support	Yes	Yes	Yes
4	2	18-39	Male	White	W	Yes	No	4-year	High	Never Heard	Willing	Support	Yes	Yes	Yes
5	2	18-39	Male	Black	W	Yes	No	4-year	High	Never Heard	Willing	Support	Yes	Yes	Yes
6	2	18-39	Female	Hispanic	W	Yes	No	≤ 2-year	High	Never Heard	Willing	Support	Yes	Yes	Yes
7	2	18-39	Male	Asian	W	Yes	No	4-year	High	Never Heard	Willing	Support	Yes	Yes	Yes
8	3	60+	Female	White	W	Yes	No	≤ 2-year	Lower	Never Heard	Not Willing	Support	Yes	No	No
9	3	60+	Male	White	W	Yes	No	≤ 2-year	Lower	Never Heard	Willing	Support	Yes	No	No
10	4	60+	Female	White	W	Yes	No	4-year	Lower	News	Willing	Support	Yes	Yes	Yes
11	5	18-39	Male	White	W	Yes	No	4-year	High	Never Heard	Willing	Oppose	Yes	Yes	Yes
12	5	40-59	Female	Black	W	Yes	No	4-year	High	Referral	Willing	Support	Yes	No	Yes
13	5	18-39	Female	Hispanic	E	No	Yes	≤ 2-year	High	Never Heard	Willing	Oppose	Yes	No	Yes
14	Supp.	18-39	NB,TG	White	W	Yes	No	4-year	High	Never Heard	Willing	Support	Yes	Yes	Yes

* Note: NB = Nonbinary, TG = Transgender, W = Western WA, E = Eastern WA

Persona Profiles

To enhance the initial persona profiles derived from cluster analysis and segmentation, qualitative analysis was conducted. This process incorporated demographic data and qualitative responses from both surveys to identify needs, preferences, and experiences of each persona. This approach allowed us to humanize the personas. Table 5 and Table 6 present the additional attributes developed through this qualitative analysis, providing a richer, more nuanced description of each persona.

Table 5. Final persona list with demographic information

Persona Number	Cluster Number	Age	Gender	Race/Ethnicity	Region	English Only	Parent	Education	Tech Readiness
1	1	40-59	Male	White	Western	Yes	Yes	4-year	High
2	1	40-59	Female	Asian	Western	No	Yes	4-year	High
3	2	18-39	Female	White	Western	Yes	No	4-year	High
4	2	40-59	Male	White	Western	Yes	No	4-year	Lower
5	2	18-39	Male	Black	Western	Yes	No	4-year	High
6	2	18-39	Female	Hispanic	Eastern	Yes	No	≤ 2-year	High
7	3	60+	Female	White	Western	Yes	No	≤ 2-year	Lower
8	3	60+	Male	White	Western	Yes	No	≤ 2-year	Lower
9	4	60+	Female	White	Western	Yes	No	4-year	High
10	5	18-39	Male	White	Eastern	Yes	Yes	4-year	High
11	5	40-59	Female	Black	Western	Yes	Yes	4-year	High

12	5	18-39	Female	Hispanic	Western	No	No	≤ 2-year	High
13	Supp.	18-39	NB	White	Western	Yes	No	4-year	High

Table 6. Final persona list with profiles and names

Persona Number	Cluster Number	Name	Age	Marital Status	Occupation
1	1	David	43	Divorced	Software Engineer
2	1	Yuki	40	Married	PT Assistant
3	2	Emily	32	Single	Social Services
4	2	Ryan	52	Single	Grocery Stocker
5	2	Desmond	37	Married	Banking
6	2	Maria	35	Married	Human Resources
7	3	Linda	66	Married	Retired
8	3	Richard	60	Divorced	Appliance Repair
9	4	Patricia	67	Married	Retired
10	5	Kyle	31	Married	Firefighter
11	5	Alicia	46	Re-married	Teacher
12	5	Isabel	21	Single	Pre-med Student
13	Supp.	Riley	26	In a relationship	Grad Student

Persona Cards

The final thirteen personas were formatted into graphic single page summaries or persona cards. [Figure 1 depicts one persona card (Personas 5 from Table 5). Quotes and drivers were also added to the personas at this stage.

[Figure 1. Persona 5: Desmond.]

Aspects of the persona cards were included based on their ability to provide a practical and user-friendly tool to bridge the gap between data and actionable insights. A detailed description of the derivation and relevance of each section is provided in [Figure 2.

[Figure 2. Sample persona with section descriptions.]

The full set of thirteen persona cards can be found in Appendix 3.

Discussion

The development of these personas leveraged a uniquely rich dataset, enabling nuanced representation of diverse population segments. The mixed-methods approach combined a rigorous data driven process with the depth of individual stories, behaviors, and needs. This synthesis achieved a deeper understanding of the dataset and allows for quickly gleanable insights, balancing the key priorities for the persona creation, which could not be accomplished with quantitative or qualitative methods alone.

The approach focused on creating personas that were simple and relatable, accurate, diverse, actionable and comprehensive. At times, decisions in the methodology had to be made to ensure

effectiveness by finding balance among these guiding principles. The aim was to define a collection of distinct individuals who could be intuitively understood. The personas reflect the breadth of demographic diversity in WA State, representing a range of characteristics rather than matching exact population proportions, ensuring actionable insights for public health informatics projects.

Cluster analysis served as a starting point, ensuring the final personas were guided by the survey data. While cluster analysis alone was the most statistically accurate process, these clusters did not establish a set of personas that upheld the other key priorities. Through a mixed-methods approach, distinct groups with more specific demographic characteristics were constructed. This technique helps highlight various user groups who might have more specific needs. For example, some of the personas can help understand challenges for those who have lower technological readiness that could benefit from targeted strategies regarding technology adoption and usage. While this is a minority of the population, individuals with these needs must still be considered in design, development, and communication.

This methodology aimed to mitigate potential biases from both human-driven and purely quantitative approaches to persona development. Traditional persona creation often relies heavily on individual interpretation which can introduce unconscious human biases and assumptions. Conversely, purely data-driven approaches can amplify existing biases in the source data which can result in further marginalizing groups underrepresented in the dataset. The mixed-methods approach employed in this study leverages the strengths of both while working to mitigate their weaknesses. The cluster analysis provided a foundation that reduced subjective biases in the initial grouping, while the qualitative analysis preserved a human perspective allowing for the creation of authentic and relatable personas that addressed gaps in the quantitative results. This approach ensures statistically grounded and representative personas.

This development process considered a wide range of demographic and outcome variables, ensuring the inclusion of diverse population segments. This comprehensive approach is crucial in public health initiatives, where efforts impact a broad population in highly personal and individual ways. While resource-intensive, this approach offers a useful framework for creating representative and realistic personas. The template may be used and adapted to meet the specific needs of various projects.

This methodology demonstrates the value of a mixed-methods approach in capturing the complexity of public health user groups and provides an adaptable template for future initiatives.

Utility of Personas in Public Health Informatics Projects

The personas developed through this process may serve as a tool to inform and guide public health informatics initiatives in the following ways:

1. **Facilitate Formative Work:** Prior to embarking on a development project, personas can be used to understand potential users and recruit for interviews and focus groups during which tailored information needs and user requirements data can be collected and analyzed.
2. **Design and Development:** The design and development process can be aided by personas providing an understanding of user preferences and behaviors. Decision making can be

aided by considering the impact of a tool or change on all target personas.

3. Usability Testing: Personas can be used to either simulate user interactions during usability testing or assist in recruitment of diverse and appropriate usability testers.
4. Communications: Personas can be used to facilitate internal communications (i.e., development team decision-making) and external communications (e.g., deciding on targets and strategies for education and information campaigns).

Implications for Public Health Initiatives

Understanding the needs and motivation behind different personas can enhance the effectiveness of public health interventions. Strategies can be tailored to specific personas, and more targeted and impactful campaigns and technologies can be implemented. A user-centered design process that is still based in comprehensive data enhances usability and ensures that interventions resonate with distinct segments of the population. This may result in higher engagement and better outcomes.

Further, as new data become available or as public health priorities shift, personas can be updated or added to reflect these changes, ensuring that the technology remains relevant and responsive to current needs. This adaptability is particularly valuable in the rapidly evolving landscape of public health, where timely and accurate information is crucial.

Moreover, the iterative and flexible nature of persona development allows for continuous refinement and adaptation. The personas developed in this study can serve as a template that other public health initiatives can modify according to their specific needs and available resources. Even without access to extensive datasets, personas can be developed using smaller datasets, targeted surveys, focus groups, or stakeholder interviews. This process and the final product can be adapted to whatever resources are available and relevant to the effort. This adaptive approach ensures that initiatives remain timely and responsive and allows for a meaningful tool across varying resource availability and project scales.

Limitations

There are several important limitations to consider regarding both the development and potential use of these personas. Most importantly, while personas are valuable tools for design and development, they are not substitutes for direct user engagement and feedback. These personas should ideally be used in conjunction with, rather than in place of, real user data and testing. They are meant to serve as practical tools when user engagement is challenging or impossible due to resource constraints, but they should not be seen as comprehensive replacements for actual user research.

Additional limitations stem from the survey data on which these personas are based. First, relying on survey data introduces sampling bias, particularly in the case of the second-phase survey data, as survey respondents may not be representative of the diverse population of WA State. Second, survey data were collected in the context of the COVID-19 pandemic, a moment in time. Public opinions rapidly evolve, and the sentiments expressed in the surveys and subsequently used to develop the personas may only reflect that context and not be applicable to a new context. Third, the surveys focused on capturing opinions regarding WA Verify and its potential expanded use cases so generalizing to other public health tools will depend on each tool

and its context.

The methodology employed presents additional limitations. Cluster analysis requires incorporating assumptions and parameter choices which may impact their interpretation, reliability, and generalization outside the context of the analysis. The development of personas may inherently oversimplify the complexity of each individual represented. In addition, bias may be introduced into the personas by including demographic and other characteristics in their development.

The addition of the Supplemental Gender Cluster, while intended to include individuals who identify as a minority gender group, presents another limitation. This supplemental cluster allowed for representation of nonbinary and transgender individuals, but it oversimplifies a diverse group that may not have similarities relevant to public health system development, use, and opinions. This limitation stems from the original data collection which had limited representation from these groups.

Future Work

The overall usefulness of these personas would benefit from an iterative process of feedback with stakeholders, including developers, designers, communication teams, and any other potential users of the personas. Stakeholders who are experienced with using personas as well as those new to using personas would provide valuable feedback. This process can help to ensure the relevance and applicability of these tools.

In particular, one question to be addressed is the appropriate level of detail that is useful for these personas. In UCD, there are differing opinions on the optimal level of detail in personas. Emerging ways of thinking promote more pared-down personas where there is no image and fewer personal details, focusing instead on only the practical elements that directly affect usability [33] [34]. Others advocate for more life-like personas with carefully considered details [35] [36]. Some details added to a persona may make the fictitious person more tangible, but some argue that these details could be irrelevant, extraneous, or distracting, potentially even introducing unnecessary bias. Future work should explore the impact of different levels of detail in personas on their effectiveness in public health informatics projects.

Additionally, more can be done to ensure the inclusiveness of personas. This can be accomplished with further user inquiry focusing on historically marginalized groups. Persona Cards could also be iteratively reviewed with groups of people whom they presume to represent to check for validity and identify any constructed biases. Without additional resources to conduct focus groups or user research, other processes can be implemented to better represent these communities such as a Backwards User Journey [30].

Conclusion

This study describes the process of developing personas from existing survey datasets. The mixed-methods approach—combining quantitative cluster analysis with qualitative data integration—resulted in 13 detailed personas. These personas may be helpful in various stages of public health technology

initiatives including needs assessment, design, roll-out and evaluation. By bridging the gap between large-scale data and actionable insights, these personas offer a powerful tool for creating more user-centered and effective public health information systems. The methodology presented here provides a flexible framework that can be adapted to meet the specific needs of various public health initiatives, potentially enhancing the use of UCD principles and thus the user-centeredness and effectiveness of public health technologies.

Acknowledgments

This work was funded through an Interagency Agreement between the Washington State Department of Health and the University of Washington School of Public Health (HED Contract 25742).

Conflicts of Interest

None declared.

Appendices

Appendix 1. Statewide Survey Instrument

The complete 32-question survey instrument used for statewide data collection.

Appendix 2. Second Phase Survey Instrument

The complete 53-question survey instrument used for the second phase of data collection.

Appendix 3. Persona Cards

The complete set of 13 persona cards developed through this research.

Abbreviations

COVID-19: Coronavirus Disease 2019

KAP: Knowledge, Attitudes, and Practices

NB: Nonbinary

TG: Transgender

UCD: User-Centered Design

UW: University of Washington

WA: Washington

WA Notify: Washington's COVID-19 Exposure Notification Tool

WA Verify: Washington's Digital Vaccine Record System

References

1. Aziz HA. A review of the role of public health informatics in healthcare. J Taibah Univ Med Sci 2017 Feb 1;12(1):78-81. doi: 10.1016/j.jtumed.2016.08.011

2. Yasnoff WA, O'Carroll PW, Koo D, Linkins RW, Kilbourne EM. Public Health Informatics: Improving and Transforming Public Health in the Information Age. *J Public Health Manag Pract* 2000;6(6):67.
3. Marsden N, Haag M. Stereotypes and Politics: Reflections on Personas. *Proc 2016 CHI Conf Hum Factors Comput Syst New York, NY, USA: Association for Computing Machinery; 2016. p. 4017–4031. doi: 10.1145/2858036.2858151*
4. Salminen J, Jung S-G, Jansen BJ. Detecting Demographic Bias in Automatically Generated Personas. *Ext Abstr 2019 CHI Conf Hum Factors Comput Syst New York, NY, USA: Association for Computing Machinery; 2019. p. 1–6. doi: 10.1145/3290607.3313034*
5. Vredenburg K, Mao J-Y, Smith PW, Carey T. A survey of user-centered design practice. *Proc SIGCHI Conf Hum Factors Comput Syst New York, NY, USA: Association for Computing Machinery; 2002. p. 471–478. doi: 10.1145/503376.503460*
6. Pruitt J, Adlin T. *The Persona Lifecycle: Keeping People in Mind Throughout Product Design*. Elsevier; 2010. ISBN:978-0-08-045573-0
7. Miaskiewicz T, Kozar KA. Personas and user-centered design: How can personas benefit product design processes? *Des Stud* 2011 Sep;32(5):417–430. doi: 10.1016/j.destud.2011.03.003
8. Holden RJ, Kulanthaivel A, Purkayastha S, Goggins KM, Kripalani S. Know thy eHealth user: Development of biopsychosocial personas from a study of older adults with heart failure. *Int J Med Inf* 2017 Dec 1;108:158–167. doi: 10.1016/j.ijmedinf.2017.10.006
9. Massey PM, Chiang SC, Rose M, Murray RM, Rockett M, Togo E, Klassen AC, Manganello JA, Leader AE. Development of Personas to Communicate Narrative-Based Information About the HPV Vaccine on Twitter. *Front Digit Health* 2021;3. Available from: <https://www.frontiersin.org/articles/10.3389/fdgth.2021.682639> [accessed Oct 6, 2023]
10. Ten Klooster I, Wentzel J, Sieverink F, Linssen G, Wesselink R, Van Gemert-Pijnen L. Personas for Better Targeted eHealth Technologies: User-Centered Design Approach. *JMIR Hum Factors* 2022 Mar 15;9(1):e24172. doi: 10.2196/24172
11. Turner AM, Reeder B, Ramey J. Scenarios, personas and user stories: User-centered evidence-based design representations of communicable disease investigations. *J Biomed Inform* 2013 Aug 1;46(4):575–584. doi: 10.1016/j.jbi.2013.04.006

12. Maguire M. Methods to support human-centred design. *Int J Hum-Comput Stud* 2001 Oct 1;55(4):587–634. doi: 10.1006/ijhc.2001.0503
13. Velsen LV, Wentzel J, Gemert-Pijnen JEV. Designing eHealth that Matters via a Multidisciplinary Requirements Development Approach. *JMIR Res Protoc* 2013 Jun 24;2(1):e2547. doi: 10.2196/resprot.2547
14. Chapman CN, Milham RP. The Personas' New Clothes: Methodological and Practical Arguments against a Popular Method. *Proc Hum Factors Ergon Soc Annu Meet SAGE Publications Inc*; 2006 Oct 1;50(5):634–636. doi: 10.1177/154193120605000503
15. Salminen J, Guan K, Jung S-G, Chowdhury SA, Jansen BJ. A Literature Review of Quantitative Persona Creation. *Proc 2020 CHI Conf Hum Factors Comput Syst New York, NY, USA: Association for Computing Machinery*; 2020. p. 1–14. doi: 10.1145/3313831.3376502
16. Jansen A, Van Mechelen M, Slegers K. Personas and Behavioral Theories: A Case Study Using Self-Determination Theory to Construct Overweight Personas. *Proc 2017 CHI Conf Hum Factors Comput Syst New York, NY, USA: Association for Computing Machinery*; 2017. p. 2127–2136. doi: 10.1145/3025453.3026003
17. Hajian S, Bonchi F, Castillo C. Algorithmic Bias: From Discrimination Discovery to Fairness-aware Data Mining. *Proc 22nd ACM SIGKDD Int Conf Knowl Discov Data Min New York, NY, USA: Association for Computing Machinery*; 2016. p. 2125–2126. doi: 10.1145/2939672.2945386
18. Siegel DA. The mystique of numbers: belief in quantitative approaches to segmentation and persona development. *CHI 10 Ext Abstr Hum Factors Comput Syst New York, NY, USA: Association for Computing Machinery*; 2010. p. 4721–4732. doi: 10.1145/1753846.1754221
19. Molino A, Revere D, Hills RA, Elder AS, West LM, Karras BT, Baumgartner C, Baseman JG. Exploring Attitudes and Obstacles around Digital Public Health Tools: Insights from Washington State's Vaccine Verification System. *JMIR Prepr* 2024 Sep 20; doi: 10.2196/preprints.66550
20. Gower JC. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika Oxford University Press*; 1966;53(3–4):325–338. doi: 10.1093/biomet/53.3-4.325
21. Ji J, Pang W, Zhou C, Han X, Wang Z. A fuzzy k-prototype clustering algorithm for mixed numeric and categorical data. *Knowl-Based Syst* 2012 Jun 1;30:129–135. doi: 10.1016/j.knosys.2012.01.006

22. Ji J, Bai T, Zhou C, Ma C, Wang Z. An improved k-prototypes clustering algorithm for mixed numeric and categorical data. *Neurocomputing* 2013 Nov 23;120:590-596. doi: 10.1016/j.neucom.2013.04.011
23. Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math* 1987 Nov 1;20:53-65. doi: 10.1016/0377-0427(87)90125-7
24. Berens E-M, Vogt D, Messer M, Hurrelmann K, Schaeffer D. Health literacy among different age groups in Germany: results of a cross-sectional survey. *BMC Public Health* 2016 Nov 9;16(1):1151. doi: 10.1186/s12889-016-3810-6
25. Gordon NP, Hornbrook MC. Differences in Access to and Preferences for Using Patient Portals and Other eHealth Technologies Based on Race, Ethnicity, and Age: A Database and Survey Study of Seniors in a Large Health Plan. *J Med Internet Res* 2016 Mar 4;18(3):e50. PMID:26944212
26. Goswami A, Dutta S. Gender Differences in Technology Usage—A Literature Review. *Open J Bus Manag Scientific Research Publishing*; 2015 Dec 2;4(1):51-59. doi: 10.4236/ojbm.2016.41006
27. Haluza D, Wernhart A. Does gender matter? Exploring perceptions regarding health technologies among employees and students at a medical university. *Int J Med Inf* 2019 Oct 1;130:103948. doi: 10.1016/j.ijmedinf.2019.08.008
28. Mantwill S, Monestel-Umaña S, Schulz PJ. The Relationship between Health Literacy and Health Disparities: A Systematic Review. *PLOS ONE Public Library of Science*; 2015 Dec 23;10(12):e0145455. doi: 10.1371/journal.pone.0145455
29. Mitchell UA, Chebli PG, Ruggiero L, Muramatsu N. The Digital Divide in Health-Related Technology Use: The Significance of Race/Ethnicity. *The Gerontologist* 2019 Jan 9;59(1):6-14. doi: 10.1093/geront/gny138
30. Spillers F. How to create Inclusive Personas. *Exp Dyn*. 2023. Available from: <https://www.experiencedynamics.com/how-to-create-inclusive-personas/> [accessed Jul 12, 2024]
31. LeRouge C, Ma J, Sneha S, Tolle K. User profiles and personas in the design and development of consumer health technologies. *Int J Med Inf* 2013 Nov 1;82(11):e251-e268. doi: 10.1016/j.ijmedinf.2011.03.006
32. Salminen J, Guan K, Jung S-G, Jansen BJ. A Survey of 15 Years of Data-Driven Persona Development. *Int J Human-Computer Interact* Taylor & Francis; 2021 Nov 8;37(18):1685-1708. doi:

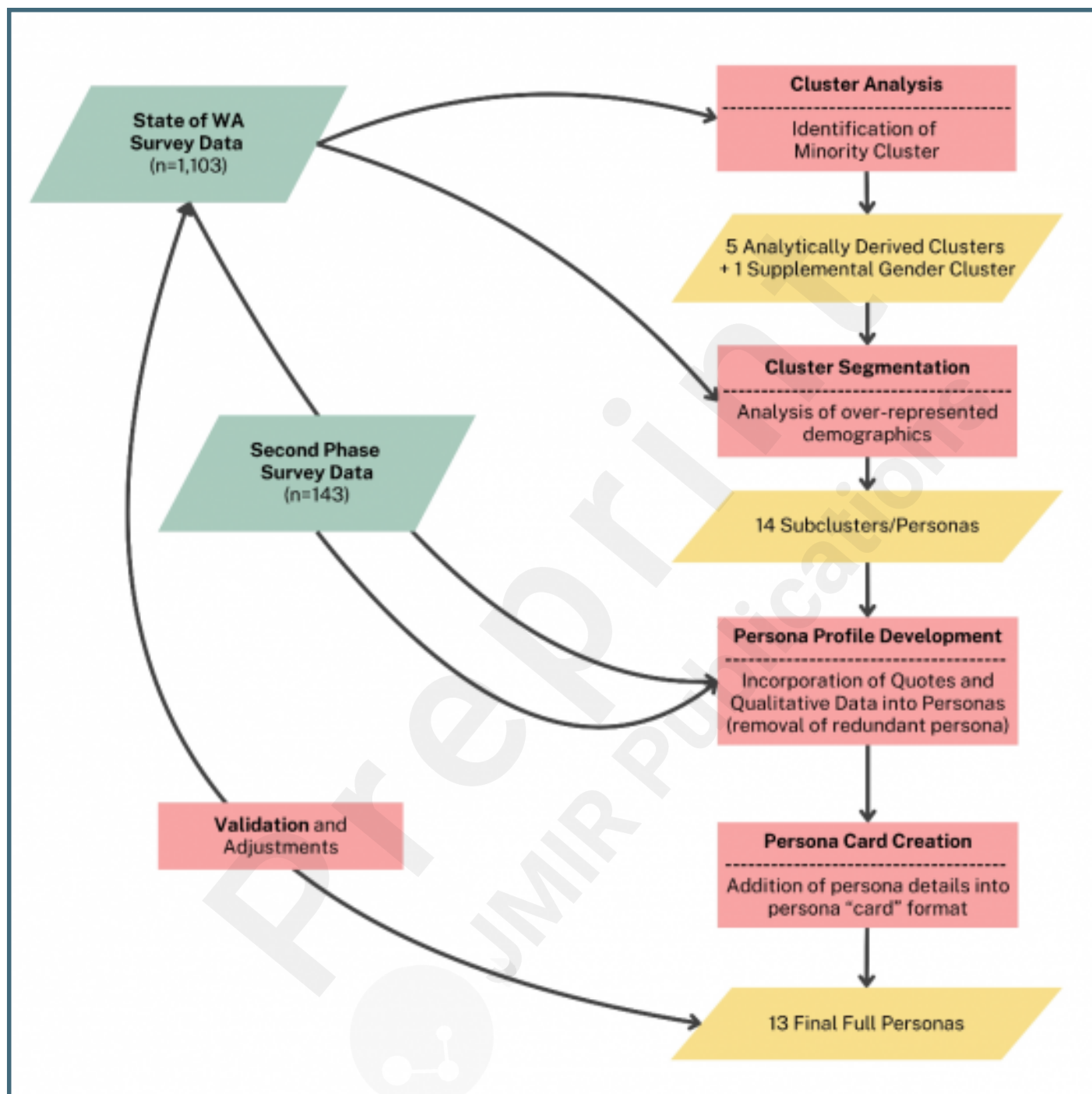
10.1080/10447318.2021.1908670

33. Malsam M. It's Time to Transition to Inclusive Personas. dscout.com. Available from: <https://dscout.com/people-nerds/transition-personas> [accessed Jul 15, 2024]
34. Polkosky M. When less is more: Minimalist personas in UX. Medium. 2020. Available from: <https://uxdesign.cc/when-less-is-more-minimalist-personas-in-ux-f2ba742f0634> [accessed Aug 18, 2024]
35. Salminen J, Nielsen L, Jung S-G, An J, Kwak H, Jansen BJ. "Is More Better?": Impact of Multiple Photos on Perception of Persona Profiles. Proc 2018 CHI Conf Hum Factors Comput Syst New York, NY, USA: Association for Computing Machinery; 2018. p. 1-13. doi: 10.1145/3173574.3173891
36. Pruitt J, Grudin J. Personas: practice and theory. Proc 2003 Conf Des User Exp New York, NY, USA: Association for Computing Machinery; 2003. p. 1-15. doi: 10.1145/997078.997089

Supplementary Files

Figures

Steps of the persona development process.



Persona 5: Desmond.

Desmond

Technically savvy and embraces the convenience technology offers

 Occupation: Banking

 Region: Western WA

 Age: 37

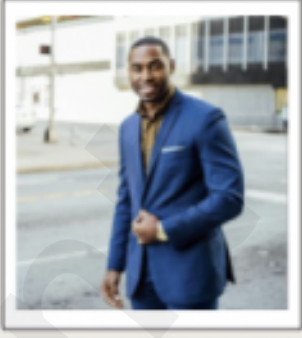
 Education: 4-year

 Race/Ethnicity: Black

 Has Minor Child: No

 English Only Home: Yes

 Gender: Male



Public Health & Tech Characteristics

- High Level of Support for PH Policies
- Owns a Smartphone
- High Tech Readiness
- Uses Online Health Tools

Drivers

- Believes technology makes everything convenient but is wary of sharing data online
- Has witnessed malicious use of stolen information through his work
- Values personal safety and uses tools that will support his own and his family's safety
- Aware of the danger of algorithms perpetuating bias

Quotes

"Data and technology can be so helpful and really help make things faster and more convenient in everyday life. The problem is that there are always people with malicious intent out there who can use any information they get their hands on to cause harm."

"I only have one dedicated credit card that I use online and I track it regularly. I am hesitant to create new accounts and share health or other personal information online. I weigh the potential risks with the benefits to make a decision about any new app or account."

"I am often out and about and since my wife has an autoimmune disease, I really value the security of knowing that there is some level of safeguard around us. To us, downloading and using online public health tools are usually worth it."

Sample persona with section descriptions.

The demographic section offers basic information about the persona. The information provides background for designers working on tailored solutions and a starting point for recruitment for interviews or focus groups.

Yuki
Enthusiastic yet cautious user of health technology

Demographic - Occupation: PT Assistant - Age: 40 - Race/Ethnicity: Asian - English Only Home: No - Region: Western US - Education: 4 year - Has Minor Child: Yes - Gender: Female	
Public Health & Tech Characteristics - High Level of Support for PH Policies - Owens a Smartphone - High Tech Enthusiast - Uses Online Health Tools	Drivers - Needs reassurance of benefits as well as how well risks have been handled before fully trusting a new digital tool - Wants to feel very confident in a system before advocating for its use by her parents - Appreciates the convergence of personal health tools
Quotes "Convenience is a big deal when it comes to health tools. It can be a game-changer for keeping us informed and safe – both personally and as a community." "I appreciate when health tools have translations so that my parents, who speak Japanese, can easily use them." "Convenience is key for me, and I find online tools helpful for tracking my and the kids' health records." "I value having easy access to my parents' medical records in Japanese."	

Drivers refer to the key motivations, needs, and factors that influence an individual's behavior, preferences, and decision-making. Including drivers in personas is crucial because they can provide deeper insights into the psychological and emotional aspects that shape a person's interactions with a product, service, or technology.

Public health and tech characteristics refer to specific traits, behaviors, and attitudes related to an individual's engagement with public health initiatives and technology. These characteristics provide insights into how a person interacts with health-related information, services, and technologies.

Quotes are excerpts or direct statements attributed to the fictional individual represented by the personas. Quotes were derived from survey responses and are intended to capture authentic voices and perspectives.

Multimedia Appendixes

Statewide survey instrument. The complete 32-question survey instrument used for statewide data collection.

URL: <http://asset.jmir.pub/assets/0dc08a06aaa877e3bf1736ccb5b2ea05.pdf>

Second phase survey instrument. The complete 53-question survey instrument used for the second phase of data collection.

URL: <http://asset.jmir.pub/assets/f141a6ace6a0df5aad5dc336cc9eec5b.pdf>

Persona cards. The complete set of 13 persona cards developed through this research.

URL: <http://asset.jmir.pub/assets/189541add51f85f29f8fccd2073d66d1.pdf>

