Collecting Tobacco Use Information with a Web-Based Family History Tool

James R Barrett, PhD; Sandra L Pelletier, PhD; Wendy Cohn, PhD; William A Knaus, MD
Department of Public Health Sciences, University of Virginia, Health System, Charlottesville VA 22908

Corresponding Author:
James R Barrett, PhD
Department of Public Health Sciences
University of Virginia
Clinical Informatics Division
Health System
Charlottesville VA 22901
USA
Email: barrett@virginia.edu

Abstract

Background: We report here on an extension to a family health history collection and interpretation tool to gather information about tobacco usage. Diagnosis and risk assessment of heritable conditions requires a family health history. We originally developed HealthHeritage, a patient-driven web-based tool to collect, update and interpret family health histories for 89 hereditary conditions, to assist health consumers or patients and their physicians [1-3]. Tobacco use is widely acknowledged to be an important, preventable environmental aspect of health [4] and an important risk factor for many hereditary conditions in HealthHeritage. Our current effort adds collection of tobacco usage by individuals and their family members to this family history tool.

Objectives: Here, we extend the original HealthHeritage to include the collection of personal and family tobacco usage history and conduct a formative evaluation of this tobacco usage data collection. This evaluation assesses comprehension and ability to answer these tobacco use questions including details important for risk assessment and research.

Methods: Based on an evidence-based literature review by medical specialists, we added tobacco use as a risk factor for 9 cancers (such as pancreatic, bladder, and kidney cancer) and 14 conditions in the overlapping areas of cardio-/neuro-vascular disease (e.g. atherosclerosis). We also developed a separate tobacco use risk assessment to address significant health risks associated with tobacco use that, although not currently shown to be strongly hereditable (as judged by the medical specialists review) have important health implications.

Our tobacco use questions resulted from a review of the literature assessing tobacco risks and several national health surveys. We chose to combine the NHLBI Atherosclerosis Risk in Communities (ARIC) survey, and the National Health Interview Survey (NHIS) as they were the most extensive in their tobacco related questions. This met an important need of one of our stakeholders, to include comprehensive data collection deemed useful for medical researchers.

We conducted a pilot evaluation of these tobacco use questions with 20 health information consumers who were patients in a Family Medicine department with a history of tobacco use. Participants were given a paper-based version of the tobacco questions; this version used screenshots of the web-based questions with instructions added to allow self-guided completion. After answering the survey participants then completed a post-survey interview. This interview clarified participant comments and questions, and went into depth about potential survey difficulties (e.g. difficulties with calculations, amount known about a relative’s tobacco use).

Results: Results include overall ease with the question content including comprehension and respondent’s ability to answer. Some difficulties were identified in certain subsets of participants (e.g. older individuals) in areas such as recall of tobacco usage dates and information about their relatives.

Conclusions: Our effort to integrate personal and familial tobacco use history into a comprehensive family history provides a tool with potential utility for in primary care and public health research. Our pilot evaluation suggests that, for these respondents, collection of personal and familial tobacco use history with detail sufficient for risk assessment and research is possible. We discuss recommendations to enhance data collection in these areas where recall was difficult. Furthermore, in our qualitative interviews participants expressed ideas that support a potential educational impact of the survey itself on participants; the implications of this aspect of the tool will also be discussed.
KEYWORDS
Tobacco; smoking; family medical history

Introduction

Results of the human genome project and related research underscore the importance of genetics for many health conditions. As genetic testing becomes increasingly available, collecting family medical histories and risk stratification based on these will be even more important in health care practices [5]. Furthermore, family health history based risk stratification can tailor disease preventative measures in public health [6,7] and in individualized care [8]. Various aspects of HealthHeritage, a prototypical family health history tool, have been previously discussed [1-3]. Important features of this web-based tool include: (1) Medical specialists guide the evidence-based development, (2) 89 strongly heritable, chronic adult onset conditions with a large public health impact in cardiac, vascular, neurological, endocrine, and oncology areas are assessed and (3) health consumers complete and update their family histories and can share the interpretation with their physicians. HealthHeritage meets most criteria identified by experts reviewing genetics in primary care and by a Center for Disease Control sponsored working group assessing family history for public health and disease prevention [9,10]. These criteria include patient completion, adapting to patient characteristics and support of prioritized clinical decision making.

Recently we added tobacco usage to HealthHeritage. Tobacco use is an important environmental risk factor [4] impacting many heritable diseases. Collection of extensive family history and relevant risk factors is critical for hereditary disease diagnosis particularly in complex, multi-factorial diseases and syndromes [11,12]. Research assessing linkage between risk factors and hereditary disease also requires this comprehensive data. Research continues to clarify the role of tobacco, including identification of a genetic component for tobacco use [13,14] and additional evidence for a genetic component to tobacco risk in the diseases it impacts [15]. In summary, the following factors compelled us to include tobacco risk in HealthHeritage 2.0, a family history tool also addressing tobacco usage:

1. Tobacco usage is an important modifiable and preventable risk factor.
2. Tobacco use impacts personal risk for many chronic adult onset conditions in HealthHeritage.
3. Collection of tobacco use history with family health history allows assessment of familial patterns of smoking’s disease implications.
4. Comprehensive data collection enables identification of the genetic components of tobacco use and of tobacco’s effects.

Adding tobacco usage to HealthHeritage, therefore assesses an important risk factor that individuals can modify and expands the tools utility to medical researchers. As a part of this extension, we consulted with medical researchers: cancer and the tools utility to medical researchers. As a part of this important risk factor that individuals can modify and expands.

Comprehensive data collection enables identification of the 4.

3.

Tobacco use impacts personal risk for many chronic adult onset conditions with a large public health impact in cardiac, vascular, neurological, endocrine, and oncology areas are assessed and (3) health consumers complete and update their family histories and can share the interpretation with their physicians. HealthHeritage meets most criteria identified by experts reviewing genetics in primary care and by a Center for Disease Control sponsored working group assessing family history for public health and disease prevention [9,10]. These criteria include patient completion, adapting to patient characteristics and support of prioritized clinical decision making.

Recently we added tobacco usage to HealthHeritage. Tobacco use is an important environmental risk factor [4] impacting many heritable diseases. Collection of extensive family history and relevant risk factors is critical for hereditary disease diagnosis particularly in complex, multi-factorial diseases and syndromes [11,12]. Research assessing linkage between risk factors and hereditary disease also requires this comprehensive data. Research continues to clarify the role of tobacco, including identification of a genetic component for tobacco use [13,14] and additional evidence for a genetic component to tobacco risk in the diseases it impacts [15]. In summary, the following factors compelled us to include tobacco risk in HealthHeritage 2.0, a family history tool also addressing tobacco usage:

1. Tobacco usage is an important modifiable and preventable risk factor.
2. Tobacco use impacts personal risk for many chronic adult onset conditions in HealthHeritage.
3. Collection of tobacco use history with family health history allows assessment of familial patterns of smoking’s disease implications.
4. Comprehensive data collection enables identification of the genetic components of tobacco use and of tobacco’s effects.

Adding tobacco usage to HealthHeritage, therefore assesses an important risk factor that individuals can modify and expands the tools utility to medical researchers. As a part of this extension, we consulted with medical researchers: cancer and the tools utility to medical researchers. As a part of this important risk factor that individuals can modify and expands.

Comprehensive data collection enables identification of the 4.

3.

Tobacco use impacts personal risk for many chronic adult onset conditions with a large public health impact in cardiac, vascular, neurological, endocrine, and oncology areas are assessed and (3) health consumers complete and update their family histories and can share the interpretation with their physicians. HealthHeritage meets most criteria identified by experts reviewing genetics in primary care and by a Center for Disease Control sponsored working group assessing family history for public health and disease prevention [9,10]. These criteria include patient completion, adapting to patient characteristics and support of prioritized clinical decision making.

Recently we added tobacco usage to HealthHeritage. Tobacco use is an important environmental risk factor [4] impacting many heritable diseases. Collection of extensive family history and relevant risk factors is critical for hereditary disease diagnosis particularly in complex, multi-factorial diseases and syndromes [11,12]. Research assessing linkage between risk factors and hereditary disease also requires this comprehensive data. Research continues to clarify the role of tobacco, including identification of a genetic component for tobacco use [13,14] and additional evidence for a genetic component to tobacco risk in the diseases it impacts [15]. In summary, the following factors compelled us to include tobacco risk in HealthHeritage 2.0, a family history tool also addressing tobacco usage:

1. Tobacco usage is an important modifiable and preventable risk factor.
2. Tobacco use impacts personal risk for many chronic adult onset conditions in HealthHeritage.
3. Collection of tobacco use history with family health history allows assessment of familial patterns of smoking’s disease implications.
4. Comprehensive data collection enables identification of the genetic components of tobacco use and of tobacco’s effects.

Adding tobacco usage to HealthHeritage, therefore assesses an important risk factor that individuals can modify and expands the tools utility to medical researchers. As a part of this extension, we consulted with medical researchers: cancer and the tools utility to medical researchers. As a part of this important risk factor that individuals can modify and expands.

Comprehensive data collection enables identification of the 4.

3.
identified as necessary to medical researchers. The ARIC study has been conducted in four US communities starting in 1987 and has included questions about smoking since 1998. The Adult Health Behaviors Section (AHB) section of the NHIS Sample Adult questionnaire contains questions related to cigarette smoking and other health behaviors, which have been in the questionnaire since 1997.

**Development of the Tobacco Use Questions**

We modified the ARIC and NHIS tobacco questions for consistent language and to maximize tobacco usage information. We used the NHIS definition of current smoker: persons having ever smoked 100 cigarettes total who currently smokes every day or some days. We incorporated these questions into the existing family history web tool. This tool takes advantage of web programmability to insure the respondent sees only relevant questions based on their previous responses, e.g. current smokers do not see questions intended for just former smokers. General screening questions about tobacco use (in self and separately for family members) answered in the affirmative lead to detailed questions for particular tobacco products, etc.; to a maximum of twenty detailed tobacco usage questions.

**Pilot Testing of the Tobacco Use Survey**

We developed a pilot survey with our tobacco questions to inform the design of the HealthHeritage web tool enhancements (i.e. a formative evaluation). For speed of implementation and flexibility of use, we adapted the web-based questionnaire into a paper-based survey. Screenshots of each web page of tobacco questions were edited to add instructions and introductory remarks. The resulting paper survey is appended to this paper (appendix 1). We told respondents that we were most interested in knowing if they understood the survey questions and felt that they could answer the questions.

We also developed a post–survey interview (appendix 2). The purpose of the interview was to clarify any respondent questions and to elicit the thought process of the respondents while responding to particular questions. We investigated potential areas of difficulty, including recollection of start and quit dates and the questions leading to quantification of smoking. As an example of how we elicit respondents thought processes, we asked respondents to review the series of questions that quantifies their smoking and to tell us about answering them as they reviewed. We probed further about any difficulties they had and what might help them respond. We also presented them with two alternative ways to collect that information and asked which they preferred.

We recruited subjects from patients in the University of Virginia Family Medicine clinic over a period of five days until we reached our pre-determined convenience sample size of 20 persons. We only recruited patients who were current or former smokers.

**Results**

We recruited 20 participants out of a total of 367 patients seen in clinic during the 5 day recruitment period. Of these 20 participants, six were male, fourteen were Caucasian and six were African-American. Two persons were excluded from the 20 because a follow-up interview could not be conducted. Of the 18 that completed the evaluation, 12 were former smokers. The total time for the survey and post-survey interview was thirty minutes. The results below represent a combination of survey responses (e.g. precision of start date) and responses to direct and open-ended questions in the post-survey interview.

**Overall Impressions of the Survey**

All respondents answered positively (“Okay”, “Easy”, “Adequate, to the point”) about their overall impression of the survey. One respondent, employed by a survey demographics firm, stated, “Excellent [it] asked me everything I knew to ask …”. Several comments indicated an educational or motivational survey impact. Five described it as interesting or helpful. One person indicated the survey did not come across as “talking down” to her but helpful; she said: “Interesting – more information. My attempts to quit made me feel good.” Two people commented - the survey was eye opening and showed them how much cumulative smoking they had, or the survey underscored the importance of quitting now. Another said it would be helpful: “helpful, helpful to families because of the danger smoking poses”.

**Did the respondents comprehend the survey?**

Respondents overwhelmingly (17 of the 18 answering this interview question) rated the overall survey as easy to understand when asked directly. On the other hand, the paper-based nature of the instrument for our formative evaluation caused some difficulties with instructions to skip questions. This resulted in a few subjects erroneously answering both current smoker and former smoker sections.

**Can respondents provide detailed information?**

When asked to rank their ability to answer the survey, most respondents (16 of 18 answering this interview question) said the survey was easy to answer. When answering questions about dates within the survey (e.g. starting or quitting tobacco usage), 12 of 18 respondents provided a precise start age (the remaining 6 responses gave age ranges of up to 6 years). Ten of twelve former smokers provide at least the year they quit. Respondents also said in the post-survey interview that answering questions about smoking rates (e.g., the number of cigarettes per day) were not difficult.

Recalling some information appeared difficult for certain subsets of smokers. For example, smokers that quit more than 10 years ago reported some difficulties recalling the date. One person said she quit about 15 years ago and then her husband said she quit in 1983 which was 22 years ago. Another person commented that his start date was difficult to recall as it was 40 years ago: “If I did this survey 20 years before, I could remember”. Infrequent smokers also had some problems with details. One infrequent smoker of pipes and cigars said “for cigars …as I haven’t ever really smoked them that much – maybe 10 cigars in a year … Pipe smoking is also very rare – how can I answer that ‘when I quit’ because I haven’t smoked my pipe in maybe 10 years?” Lastly, elderly persons noted problems as well. One stated, “Most people my age, 77, their memory problems make the survey difficult to answer. I have Alzheimer’s …”. 
Do respondents have a consistent concept of “regular smoking”?  

Many of the published surveys we reviewed, including both NHIS and ARIC, use a phrase like “regular smoking”. For example, NHIS asks, “How old were you when you FIRST started to smoke fairly regularly?” and ARIC asks “How old were you when you first started regular cigarette smoking?” For both of these interviewer-facilitated surveys, interviewers are instructed to say that regular is ‘whatever it means to you’ if respondents ask. To avoid any risk factor misclassification due to vague meanings [21,22], we sought to determine what our respondents consider by “regular smoking”. We showed our survey question, “How old were you when you first started regular cigarette smoking?” to participants again during the post-survey interview and asked them what ‘regular’ meant to them. Most respondents - 82% (14 of 17 responses) - included “daily” in their definition. The other three answers either said that the concept includes the idea of addiction or that “regular” begins when you buy your own cigarettes.

Can respondents provide details about total life time exposure?  

Many surveys (including both NHIS and ARIC) ask respondents about long periods of not smoking in order to exclude those periods from the calculation of total life-time tobacco exposure. Only one of our seven current smokers ever quit for longer than 1 year. Most current and former smokers said the total number of years not smoked was between one and one and a half years; the remaining responses were uncertain or vague. Therefore, our preliminary results suggest that there is little value for most smokers in recalling and calculating periods of non-smoking.

Our method of assessing total life time exposure relied on a simple frequency calculation, which required respondents to estimate frequency during “the entire time [they] smoked”. We anticipated the difficulty of this question and reviewed it with participants while assessing two alternative designs based on recommendations by health outcomes and evaluation experts. Alternative 1 asked exact recall of most current frequency and the length of time at that level. Alternative 2 employed a visual representation of the averaging process. In the post-survey interviews four out of six respondents felt the current style of frequency estimation was the same or easier to answer than either alternative.

Can respondents answer about their relatives?  

Tobacco use history of relatives is critical for researching familial patterns and genetics of tobacco use. Therefore, it was important to evaluate whether respondents could provide this information. We explored this issue by asking participants to rank the likelihood that they could provide this information about their relatives. Twelve people answered: of these, six said “very likely”, three said “somewhat likely”, and three said “unlikely”. When asked the reasons for difficulties, five said it was due to deceased relatives, three said distance in relationship, and two suggested distance in location of relatives. One respondent included both spatial and relational distance: “for those relatives I know I could do this, for those relatives that are close by”. Two respondents brought up privacy issues, e.g. one stated; “rather not get into relatives … it would have to be their own decision; I wouldn’t get that information”.

We probed specific alternatives for providing dates – (1) the current style: day, month, year with flexibility to provide month, year or year only, (2) 5 year increments, (3) 10 year increments, or (4) life milestones: ages that reference life events such as high school, first job, when married. Nine of the fourteen answering said they could provide the information about more of their relatives using the 5 year increment format and five said that using life milestones would help them recall.

What if respondents are uncertain?  

We discussed with our respondents how they prefer to answer when they are uncertain and provided them with the following options: (a) leave it blank, (b) give partial answers, estimates or guesses, (c) prefer not to answer, or, (d) don’t know. Eleven out of fourteen respondents would use a partial answer, estimate or guess. Seven would use these responses for dates or rates of smoking and four of these seven would use them for personal dates or rates. Two people that would use partial answers had illustrative comments: one said, “maybe with dates but as honest as possible”, the other said, “not on myself. Mom or fiancé, would estimate length of time smoked, know the quantity and times quit.” Answering questions about relatives (particularly deceased relatives) generated a different preference; most (4) preferred ‘don’t know’ rather than either leave blank or ‘prefer not to answer’. Lastly, when understanding the question was an issue, one respondent would select ‘prefer not to answer’ and another preferred leaving a blank.

Discussion  

Most respondents provided the necessary information about their own dates and rates of smoking and most provided detailed information. Consistent with the literature regarding estimates of frequent and regular behaviors [23], averaging over long periods of time did not pose much difficulty for our respondents; difficulties mostly arose when tobacco usage was infrequent and irregular (such as occasional pipe or cigar smoking). Most respondents stated they could also (somewhat easily) provide information about relatives. However, detailed tobacco usage of relatives becomes problematic if the relative is deceased or is distant (in relationship or in physical location) and privacy concerns arose for some (two out of eighteen in our study).

Another particular problem arose when recalling prolonged periods of not using tobacco. Most respondents never quit longer than one year and for those that had most didn’t pause in their smoking habit cumulatively longer than 1.5 years. In our post-survey interview this question was pointed to as being difficult for some. Therefore, although this parameter is a component of total lifetime exposure, a very critical variable in risk assessment and research, our pilot data suggest that exclusion of this question from the survey would eliminate a difficult retrieval problem with little loss of data for total lifetime exposure. Further exploration of this with a larger group of respondents is warranted.

We suspect there is a “learning curve” as a respondent proceeds through the survey. In particular, the amount of information
about dates (or about rates of tobacco usage) became more detailed by the end of the survey for at least 3 respondents. There are several methods to measure this effect and use it to maximize data collection. Respondents could be prompted to think about their tobacco history during the introduction of the survey; this might allow respondents to provide detailed answers throughout the survey. Repeating one or two questions about important parameters (e.g., start date or smoking rate) might generate more detail in the later responses. Finally, respondents could be encouraged to return to the survey at a later time and this could be easily facilitated by the web-based format of the survey. The latter two methods would help measure intra-observer variability.

The use of life milestones, events or event history calendars may also facilitate the recall of medical histories and health behaviors. This approach uses personal experiences (or in some cases cultural events) as cues to recall or to prevent some recall errors [24,25]. Five of our respondents felt that life events could aid recall of relative dates and in fact many referred to such events when responding during the interview (e.g. when in college, after the divorce, after my heart attack). This technique seems well suited to family history collection and it may provide a method of prompting recall of personal and family health and tobacco history details.

Based on respondent comments, aspects of this tool might also provide an opportunity for education. That is, some respondents suggested they learned things about their own tobacco usage habits, such as the importance of cumulative years of smoking. Others commented that the tool provided positive feedback about their attempts to quit. Some of the respondents felt that this feedback would also be of value for their relatives. Surveys like this could augment other educational approaches related to health behaviors.

We believe HealthHeritage is unique among tobacco use tools by collecting information not only about an individual’s tobacco use, but also the use by their family (i.e. 1st and 2nd degree relatives) and doing so in the context of a comprehensive family health history. The ability to update the family health history information and share it with family members and physicians provide additional value for data collection and utility. This enhanced HealthHeritage should also be useful as a research tool for diseases in that have both a tobacco component and a strong heritable component. Further collection of comprehensive data could provide evidence for the hereditary nature of the smoking risk factor in one or more of these disease areas and identifying families for research to identify the associated genes. Finally, HealthHeritage can be enhanced in a modular fashion reflecting new research as demonstrated by the current expansion. Therefore, we expect to be able to add other risk factors, diseases and associated risk assessments into Health Heritage including any associated with tobacco usage.

Acknowledgments
The authors thank Dr Meisfeldt, Dr Tucker and Dr Worrall for their medical expertise which guided the development of the risk assessment enhancements based on tobacco use to HealthHeritage. In addition, we thank Dominion Digital and Anthony Wensel for the web-site development. This study’s protocol was approved under HIC # 2005-0064-00.

Conflicts of Interest
The authors have no conflicts of interest to declare.

Multimedia Appendix 1
Microsoft Word file- survey for Tobacco users
[WinWord (.doc) file, XX KB - somedoc.doc ]

Multimedia Appendix 2
Microsoft Word file- interview form (after taking the survey)
[WinWord (.doc) file, XX KB - somedoc.doc ]

References


17. ARIC. Home page. (The Collaborative Studies Coordinating Center (CSCC), University of North Carolina) URL: http://www.cscc.unc.edu/ARIC/index.htm [accessed 2006 31 May]


