IMPACT: An Object-oriented Graphical Environment for Construction of Multimedia Patient Interviewing Software

Leslie A. Lenert, David Michelson, Chris Flowers and Merlynn R. Bergen
Department of Medicine, Stanford Univ. School of Medicine, Stanford, CA 94305

Abstract

This paper describes our software for rapid construction of multimedia computer interviews. The program, which we call IMPACT, was designed to measure preferences for health outcomes using the standard gamble and other decision analytic techniques. IMPACT is also a multimedia shell program that allows researchers to interactively construct patient interviewing instruments without programming or scripting. It supports the integration of text, graphics, synthesized speech, digital sound and QuickTime movies into interviewing instruments through a point-and-click interface. IMPACT also supports branching logic and randomizing the presentation order of materials within an instrument. This allows customization of the presentation based on patient responses and facilitates experimental designs. Validation studies show that preference assessments performed using IMPACT have high test-retest reliability (r=0.83, n=96). Post-test surveys (n=52) show that most subjects understand valuation methods (86%) and believe that the explanations provided were clear (96%) and that methods were reasonable (80%). The majority of subjects thought the preference assessment methods were not difficult to use (53%) and would have been comfortable using such methods for medical decisions (53%).

Introduction and Background

Computer interviews have long been used to collect data about a patient's health history [1, 2], risk factors for cancer and other diseases, and preferences for medical treatment [3]. Patient interviewing software has typically has used text-based interfaces which display computerized questionnaires [1, 2]. These interviews rely on subjects to read materials on the computer screen and select appropriate responses. This limitation hampers the application of computer interviewing methods in groups with poor reading skills and lower educational levels. It also may limit the complexity of questions that can be asked in automated computer interviews. In our research to use computers to measure patient preferences for health conditions (an important task in medical decision making [4] and health policy research [5]), we initially used a combination of displayed text and animated graphics to elicit preference s. Other investigators continue to use this approach [6]. However, we believed there were important advantages in using multimedia methods to explain both the effects of a health condition on quality of life and the complex tasks required for decision analytic methods.

Encouraged by the videodiscs developed by the Foundation for Informed Patient Decision Making (which used multimedia methods to explain treatment options to patients [3]), we began research efforts aimed at using similar methods to conduct automated patient interviews for preference elicitations. Our goal was to develop software that would completely automate the process of describing and assessing preferences for health states. This would have a number of advantages, including the removal of many logistical obstacles to preference measures in clinical trials and eliminating potential bias from interviewers.

To achieve this goal, we developed a series of custom–built multimedia software programs that used voice recordings, text, pictures and digital video to communicate the effects of health conditions on quality of life. We also developed methods to explain and elicit preferences for these conditions. We employed these programs in several clinical studies where they were well accepted by patients and produced reliable and valid utility assessments [7,8]. We also demonstrated that multimedia methods improved subjects' learning about health states and reduced their uncertainty about what life might be like in those health states [9].

One lingering obstacle to computer based preference elicitation methods in Health Services Research and patient decision support has been the difficulty in constructing instruments. This paper describes our research to simplify the program development process by applying object-oriented programming principals. The software resulting from this research is called IMPACT (for Interactive Multimedia Preference Assessment Construction Tool).

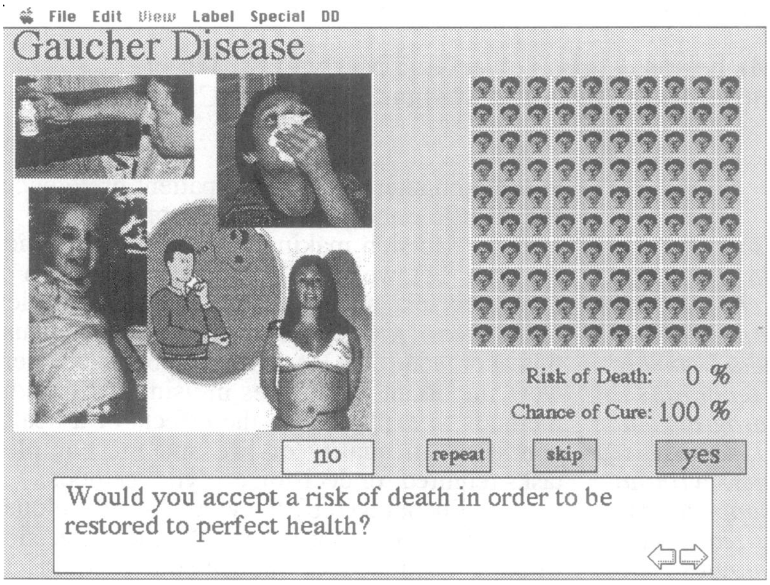

Figure 1. Scene from a computer preference assessment instrument. In this scene, a subject's preferences are being assessed using the standard gamble. At this point in the scene, the computer is asking the subject if he or she would accept any risk of death in order not to have to live with Gaucher disease (which has been described in detail in a preceding QuickTime movie.)

Design considerations

Our goal in designing IMPACT was to build a multimedia development tool aimed specifically at facilitating the construction of computer-based patient interviewing instruments. To minimize the programming effort required to develop new instruments, IMPACT reuses as many elements as possible from previously developed protocols.

We envisioned two potential classes of IMPACT researchers: disease area specialists and methodologists. Disease area specialists are researchers who are primarily interested in using validated protocols to measure preferences for a research project. This group of users would access a previously written protocol within IMPACT and make changes in the protocol to customize it for the disease area of their research. This might include changing pictures, movies and other revisions of the program's content to reflect the context of the disease. Our goal was to make this customization task easy to perform, even for researchers without computer programming skills

The second group of users we envisioned were researchers interested in preference assessment methods. They would use the tools provided by IMPACT to create new protocols or design new types of instruments. These researchers would have to be somewhat skilled at computer programming. To support these users, IMPACT would allow the assembly of new protocols from elements of previous protocols. These elements might perform general tasks, such as presenting a multiple choice question or a health state description. IMPACT would also provide support for such experimental methods as randomizing the order of protocol elements, branching logic, and support for tracking variables in the prgram and recording of data.

Implementation

We developed IMPACT using SuperCard 1.7 for Macintosh computers. IMPACT requires 12 megabytes of RAM and System 7.1 or 7.5 with QuickTime and PlainTalk system extensions. The program runs as a SuperCard stand-alone application.

IMPACT's design is shown in Figure 2. IMPACT has two parts: a *Player* and an interactive *Editor*. Both are integrated into a single program. When subjects view a program created in IMPACT, they see an interactive multimedia presentation delivered by the *Player* portion of the program. The program is run in display–only mode by the *Player*. Subjects cannot modify the program or see editing tools. Figure 1 shows an example of a scene being played by IMPACT. The *Player* components of IMPACT consist of a *Scene player*, a run-time *controller*, and two internal data files. One internal data file contains the sequence of scenes to be displayed (the *Scene Sequence File*), and the other contains instructions for how to display each *Scene* (the *Scene Data File*). The *Controller* uses the *Scene Sequence* file to determine which *Scenes* are in the protocol to display. It sends the appropriate data for each *Scene* to the *Scene Player*. The *Scene Player* displays this data, collects the responses and passes them back to the *controller*, which records the data in an external file. The *Controller* then selects the next *Scene* to be played and repeats the process until all the *Scenes* in the sequence file have been played.

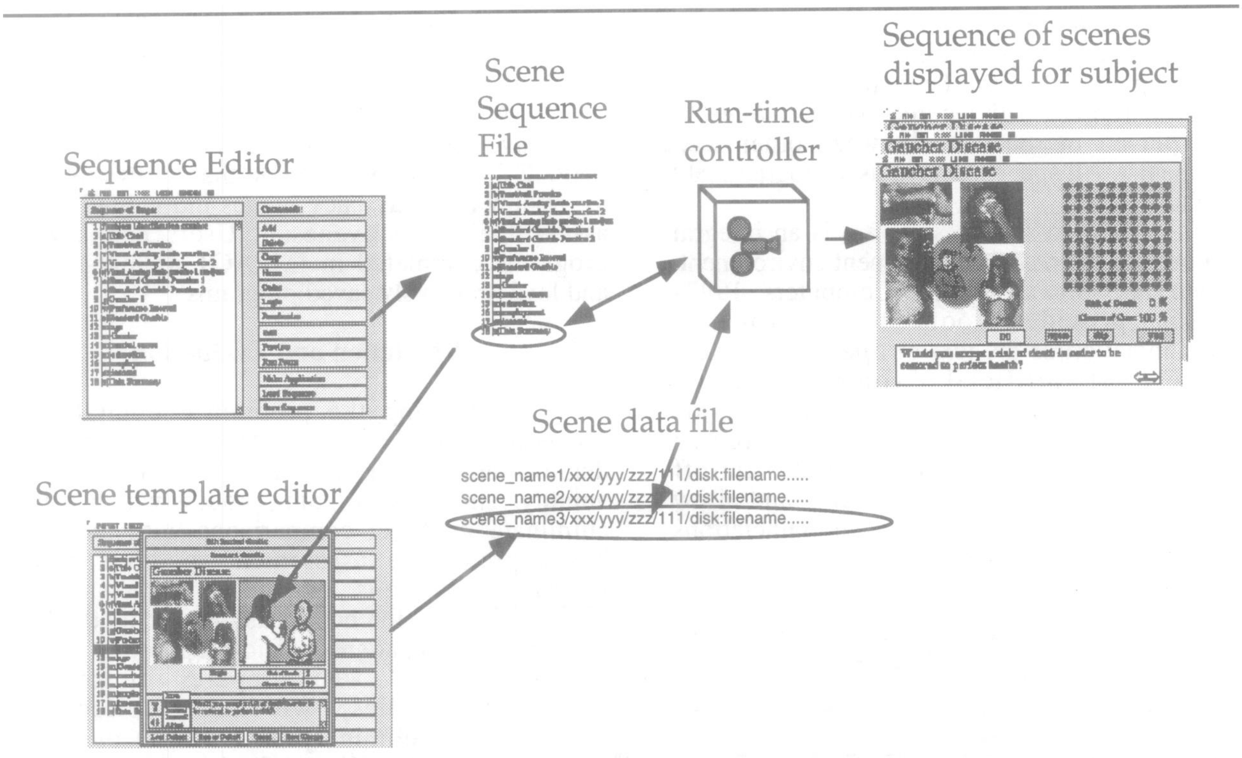

Figure 2. Schematic of the IMPACT design. Programmers use the Sequence Editor to specify the number and order of scenes in a patient interviewing instrument. They then use the Scene Template Editor to specify the details of each Scene and the types of data to be recorded from the subject. When a completed instrument is played, a run-time controller displays each Scene in the protocol one at a time by selecting the appropriate Scene from the Scene Sequence File and passing the data for that Scene to the Scene Player, which then displays the information and collects data from the subject.

Researchers access the *Editor* through a command-key function. To develop a new protocol within IMPACT, a user first chooses the scenes to include in an instrument from a list of ready–made scene templates. The types of *Scene Templates* available reflect the prototypical elements of a patient interview and include an introduction to the program; mouse training; a movie stage; free-response questions; multiple choice questions; a pair–wise comparison; standard gamble, time-trade-off utility and visual analog scale utility assessment; and a review of previous responses. New *Scene Templates* can be readily added to extend the program.

An important point of IMPACT is that users can specify conditions for the *controller* to play a given Scene. This allows the implementation of instruments to randomize a series of completed scenes or to make the order of scenes dependent on the prior response of the subject. These features are designed to improve the validity of preference assessment experiments and to allow implementation of complex experimental designs (such as fractional factorial designs) to measure preferences for a series of related health conditions [10].

After specifying the sequence of *Scenes*, a researcher uses the *Scene Editor* to program the content to be displayed in a given scene. Each scene is represented by a template. Users fill in the template to customize the scene for their applications. Different elements of scenes include *Speeches*, Movies-in-a-Window or Pictures-in-a-Window, and illustrations of risks or trade-offs used in preference elicitation. Every element of a scene has default content materials. The programming of a scene requires only specifying changes from default explanations and graphics for that scene. In addition, the *Scene Template* itself can be modified using a simple scripting language.

An example *scene template* is shown in Figure 3. Using this template, authors set text, pictures and QuickTime movies for the scene. In each *Speech*, text is displayed on the screen and synchronized with either a recorded voice or synthesized speech. The scene template in Figure 3 has five *Speeches (Intro, Choice,* *lincrease1, Increase2 and Adjust).* Users select a *Speech* to edit by clicking on its name. Users then type in the text to be displayed and spoken or synchronized with recorded voices. *Speeches* are played in sequence between other scene elements. All speeches have user programmable default values.

Text-to-speech software is an integral part of the IMPACT development environment. PlainTalk software (Apple Computers, 1992-1995) provides tools to control the pronunciation of the text of *Speeches* using a variety of artificial "voices". These mechanisms have been supplemented inside IMPACT with software to synchronize displayed text with spoken words, cut-off speeches in progress on command, and allow programmers to correct the mispronunciations of Text-to-Speech routines.

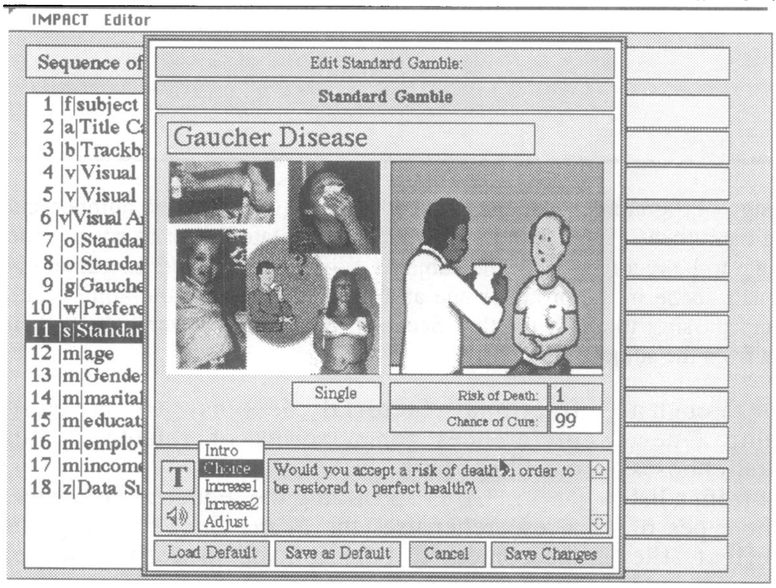

Figure 3. IMPACT Scene Template Editor. This figure shows how the instrument designer fills in the details of a particular scene in the assessment protocol. Researchers use this template to specify the title, pictures displayed and instructions to be spoken in each scene.

IMPACT has special features to support the use of utility theory to measure patient preferences. Users can precisely control the procedures for utility assessment. They can set the duration of survival for time trade-off utility assessment or set the program to look up appropriate values from a table. Users can also specify the search pattern used to find the subject's indifference point. The graphical displays used to provide visual feedback for the standard gamble and time trade-off are modular and interchangeable. New designs for displaying trade-offs or risks can be easily substituted for default displays.

After completing a *Scene Template*, users save their changes to the *Scene Data File*. IMPACT has a preview mode that allows users to review and play back scenes individually to assess their effectiveness. Users can save programs completed in IMPACT as text-files and load previously saved programs.

Validation and Evaluation

Validating the preference elicitation procedures used within IMPACT requires demonstration of their reliability and construct validity. This is an on-going process which will culminate in direct comparisons with paper-based preference assessment methods and formal evaluations of the usefulness of preference measurement methods in patient decision support.

To prepare for these studies, we have collected preliminary data on the overall reliability of the preference elicitation techniques implemented in IMPACT. The reliability of preference assessment has been examined in results of interviews of normal subjects (n=97) on two occasions one-week apart. Interviews were conducted for preferences for health states related to Gaucher disease, a rare inherited glycogen storage disorder. The 95 percent confidence interval for the absolute difference in preference ratings between two ratings perforemd one-week apart was 0.054 to 0.032. The correlation between test and retest preferences was 0.83. This compares favorably to the paper-based utility assessment methods, where correlations range for 0.54 to 0.88 [10].

In a separate study, we examined the validity of and potential role of IMPACT in medical decision support. We administered a questionnaire to a convenience sample of 51 normal subjects who had just completed of an automated computer interview. The comprehension of the standard gamble was assessed by analyzing the subject's responses to an open-ended question about the meaning of the procedure.

Table 1. Investigator-rated understanding and subject-rated clarity, difficulty, reasonableness, and comfort with preference assessment procedures used within IMPACT.

Percent rating (52 subjects)	Understanding of the standard gamble	Clarity of explanations of required tasks	Difficulty level of the rating tasks	Reasonableness of the rating tasks	Comfort with use in medical decisions
High (4)	86%	82%	14%	49%	20%
(3)	–	14%	33%	29%	33%
(2)	–	0%	33%	18%	39%
Low (1)	13%	4%	20%	4%	8%

Subjects also rated the clarity, difficulty, reasonableness of the rating task, and their comfort with using the rating tasks in medical decision making on 4 point discrete Likert-type scales.

Results of this study are shown in Table 1. A large fraction of subjects understood the concepts underlying standard gamble. Subjects thought the explanations of rating tasks were clear, but some found the rating tasks difficult. Most subjects believed the rating tasks were reasonable decisions. However, some subjects would have been uncomfortable using these methods in medical decisions making. The results suggest that preference elicitations performed using IMPACT have construct validity and that further evaluation of the program's usefulness as a clinical decision support tool is warranted.

Future Work

Future work with IMPACT will focus on additional validation of the preference elicitation procedures implemented in the program. One important aspect of validity is the internal consistency of preference elicitations obtained from a subject. We are developing methods to assess the internal consistency of valuations across different methods and within a given method of preference assessment. We are also developing interface designs that will allow us to apply these tests in an interactive fashion during computer interviews. We hypothesize that providing users with feedback on the internal consistency of their preference ratings will further improve the validity and reliability of preference assessments.

Acknowledgments

This work was supported by grants from the National Library of Medicine (LM 05626) and the Agency for Health Care and Policy Research (HS078180-02). Computer resources wereprovided by the Stanford University CAMIS project funded by the National Library of Medicine (LM-05305). We would like to acknowledge the intellectual contributions of Alan Garber and Mary Goldstein to the procedures we have developed for preference assessment.

References

1. Locke SE, Kowaloff HB, Hoff RG, et al. Computer-based interview for screening blood donors for risk of HIV. JAMA. 1992;268(10):1301-5.
2. Slack WV, Slack CW. Patient-computer dialogue. N Engl J Med 1972;286(24):1304-9.
3. Kasper JF, Mulley AJ, Wennberg JE. Developing shared decision-making programs to improve the quality of health care. Qrb 1992;18(6):183-90.
4. O'Meara JJ, McNutt RA, Evans AT, Moore SW, Downs SM. A decision analysis of streptokinase plus heparin as compared to heparin alone for deep-vein thrombosis. New Eng J Med 1994;330(26):1864-9.
5. Torrance GW. Measurement of health state utilities. J. of Health Econ. 1986;5:1-30.
6. Sumner W, Nease R, Littenberg R. U-titer: A utility assessment tool. Proc. of 15th Ann Symp on Comp App. in Med Care. McGraw-Hill, 1991: 701-705.
7. Morss S, Lenert L, Faustman W. The side effects of antipsychotic drugs and patients' quality of life: Proc. of 17th Ann. Symp. Comput Appl Med Care. McGraw-Hill, 1993: 17-21.9.
8. Goldstein MK, Michelson D, Clarke AE, Lenert LA. A multimedia preference-assessment tool for functional outcomes. Proc of the 17th Annu Symp Comput Appl Med Care. 1993;844-8.
9. Goldstein MK, Clarke AE, Michelson D, Garber AM, Bergen MR, Lenert LA. Developing and Testing a Multimedia Presentation of a Health-State Description. Med Decis Making 1994;14:336-344.
10. Froberg D, Kane R. Methodology for Measuring Health-State Preferences - II: Scaling Methods. J of Clin Epi 1989;42(5):459-471.