User Study of a Spanish-language ClinicalTrials.gov Prototype System

Graciela Rosemblat, Tony Tse
Lister Hill National Center for Biomedical Communications
National Library of Medicine, NIH, DHHS, Bethesda, MD

ABSTRACT

We conducted a user study of monolingual and bilingual Spanish-speaking consumers (n=36) to evaluate a Spanish-language ClinicalTrials.gov prototype. The prototype leverages an existing English-only consumer health resource by combining (1) Spanish-English cross-language information retrieval (CLIR) and (2) English-Spanish document display techniques. We collected user feedback on expectations, usability, and satisfaction. Preliminary results suggest improved online information access by Spanish-speakers. The goal is to develop a general approach for other systems and languages.

INTRODUCTION

Nearly two-thirds of all online English-speaking American adults have reported seeking health information [1, 2], compared to 51% of Hispanic Internet users in the US (or roughly 25% of the total US Hispanic population [3]). More Spanish speakers would likely seek health information online if additional resources were available in their native language. However, creating and maintaining such "fully translated" sites have been reported to be resource intensive [4].

Cross-language information retrieval (CLIR) is one way to leverage existing English-language consumer health information systems. While CLIR supports the submission of queries in Spanish, relevant documents will be displayed in English unless mechanisms for Spanish display are developed to supplement CLIR.

This paper reports on a user study to (1) determine end users' expectations for a Spanish-language ClinicalTrials.gov and (2) assess the usability of a Spanish-English CLIR prototype. For the individual trials, user feedback on a document surrogate ("doclet") that displays basic information about the retrieved documents in Spanish was also collected.

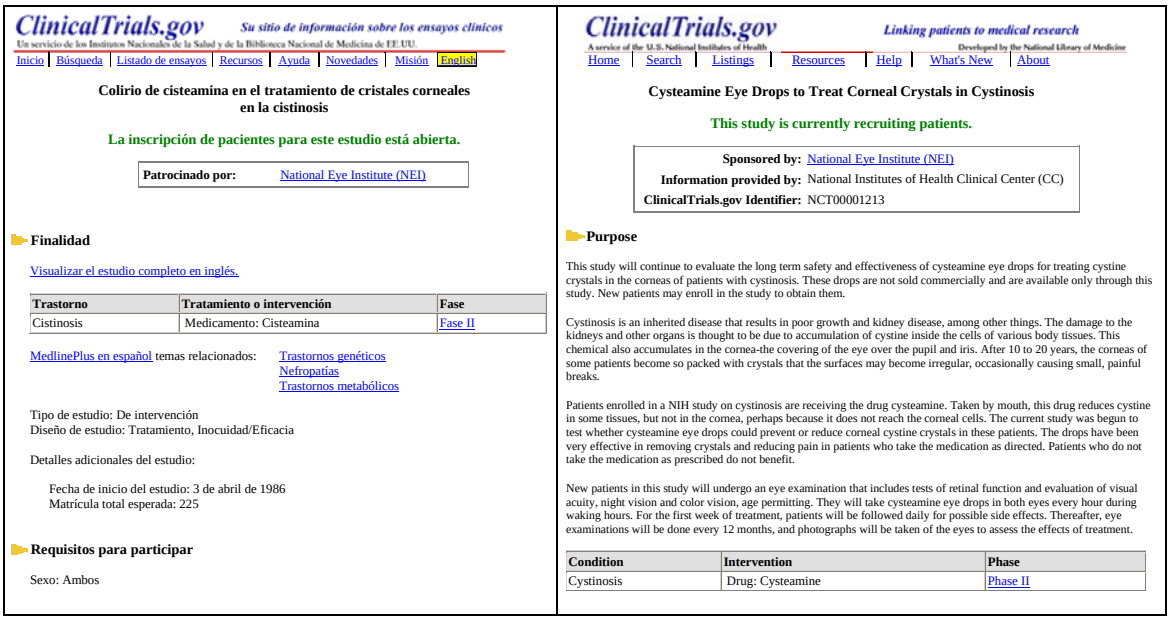
BACKGROUND

We first designed a functional Spanish-language prototype of ClinicalTrials.gov, a clinical trials registry developed by the National Library of Medicine (NLM). The dynamically generated records (32,000 as of July 2006) consist of both structured data containing key descriptors (e.g., Intervention, Condition) and free-text fields with changeable data. In contrast, static pages are used for content that changes less frequently (e.g., Help). The Spanish system consists of two distinct components: (1) translation of query terms from Spanish into English and document retrieval using CLIR, and (2) translation and display of key information in Spanish.

Our CLIR prototype matched Spanish queries with English terms via an NLM-curated bilingual term list. The CLIR parameters were fine-tuned using native English- and Spanish-language queries against an English corpus [5]. The Spanish site display contained manually translated static pages (Figure 1). Abbreviated Spanish-language clinical trial records or "doclets" (Figure 2) were created semi-automatically [6]: one-time manual translation of controlled vocabulary items for key descriptor fields allowed display in multiple doclets. In the free-text fields, where simple term look-up of changeable data was not viable, links to the corresponding English record were included for bilingual users. However, to provide some context, the descriptive Brief Title was machine-translated and manually post-edited. Proper names and locations were not translated.

Figure 1. Prototype homepage. [6]

The design of the doclet information surrogate was intended to provide users with just enough details to enable selection of studies for further consideration (e.g., with a bilingual healthcare provider). An earlier review of the overall prototype by Spanish-speaking health professionals worldwide provided favorable impressions and constructive feedback [6]. However, it was suggested that information in Spanish about the trial purpose would make the doclets more useful.



ClinicalTrials.gov — Su sitio de información sobre los ensayos clínicos
Un servicio de los Institutos Nacionales de la Salud y de la Biblioteca Nacional de Medicina de EE.UU.
Inicio | Búsqueda | Listado de ensayos | Recursos | Ayuda | Novedades | Misión | English

Colirio de cisteamina en el tratamiento de cristales corneales en la cistinosis

La inscripción de pacientes para este estudio está abierta.

Patrocinado por: National Eye Institute (NEI)

Finalidad

Visualizar el estudio completo en inglés.

Trastorno	Tratamiento o intervención	Fase
Cistinosis	Medicamento: Cisteamina	Fase II

MedlinePlus en español temas relacionados: Trastornos genéticos, Nefropatías, Trastornos metabólicos

Tipo de estudio: De intervención
Diseño de estudio: Tratamiento, Inocuidad/Eficacia

Detalles adicionales del estudio:

Fecha de inicio del estudio: 3 de abril de 1986
Matrícula total esperada: 225

Requisitos para participar

Sexo: Ambos

ClinicalTrials.gov — Linking patients to medical research
A service of the U.S. National Institutes of Health — Developed by the National Library of Medicine
Home | Search | Listings | Resources | Help | What's New | About

Cysteamine Eye Drops to Treat Corneal Crystals in Cystinosis

This study is currently recruiting patients.

Sponsored by: National Eye Institute (NEI)
Information provided by: National Institutes of Health Clinical Center (CC)
ClinicalTrials.gov Identifier: NCT00001213

Purpose

This study will continue to evaluate the long term safety and effectiveness of cysteamine eye drops for treating cystine crystals in the corneas of patients with cystinosis. These drops are not sold commercially and are available only through this study. New patients may enroll in the study to obtain them.

Cystinosis is an inherited disease that results in poor growth and kidney disease, among other things. The damage to the kidneys and other organs is thought to be due to accumulation of cystine inside the cells of various body tissues. This chemical also accumulates in the cornea-the covering of the eye over the pupil and iris. After 10 to 20 years, the corneas of some patients become so packed with crystals that the surfaces may become irregular, occasionally causing small, painful breaks.

Patients enrolled in a NIH study on cystinosis are receiving the drug cysteamine. Taken by mouth, this drug reduces cystine in some tissues, but not in the cornea, perhaps because it does not reach the corneal cells. The current study was begun to test whether cysteamine eye drops could prevent or reduce corneal cystine crystals in these patients. The drops have been very effective in removing crystals and reducing pain in patients who take the medication as directed. Patients who do not take the medication as prescribed do not benefit.

New patients in this study will undergo an eye examination that includes tests of retinal function and evaluation of visual acuity, night vision and color vision, age permitting. They will take cysteamine eye drops in both eyes every hour during waking hours. For the first week of treatment, patients will be followed daily for possible side effects. Thereafter, eye examinations will be done every 12 months, and photographs will be taken of the eyes to assess the effects of treatment.

Condition	Intervention	Phase
Cystinosis	Drug: Cysteamine	Phase II

Figure 2. Doclet (left) and corresponding English-language document (right). [5]

This feedback prompted us to look into an alternative doclet display format, in which a translated excerpt of the English purpose description could better facilitate user decision-making and relevance judgments:

> The optimal amount of information needed in a given decision-making situation lies somewhere along a continuum from "not enough" to "too much" [7]

We thus developed an alternative "doclet+purpose" design for evaluation that included a brief Spanish-language "purpose statement:" up to three sentences algorithmically extracted from the English Purpose section, machine-translated into Spanish, and manually post-edited for inclusion in the doclet. In Figure 2, this extracted text would include only the first paragraph of the Purpose (document on right).

The objective of the user evaluation reported in this paper was to observe the interaction of actual mono- and bilingual Spanish-speaking consumers with components of the prototype design, and obtain feedback on how the prototype and doclet met these expectations. In particular, the goals were to:

(1) Examine end users' information needs;
(2) Determine users' abilities to obtain relevant information using the prototype; and
(3) Assess user satisfaction with the doclets, compared to the original English-language records and the "doclet+purpose."

METHODOLOGY

Study Design

The study consisted of three parts (referred to as "parts 1-3" in this paper), one for each goal:

(1) Understanding clinical research and expectations for a Spanish clinical trials information resource;
(2) Usability of the prototype Website; and
(3) Satisfaction with the Spanish-language doclets.

Each 45-60-minute session was video recorded. Screen capture and recording software (Morae™ from TechSmith) was used in the hands-on usability testing (part 2) for later analysis of user actions. Spanish was used in all parts and aspects of the study.

Participant Selection: A total of 36 (26 female, 10 male) monolingual and bilingual US-based Spanish-speaking consumers with some online information seeking experience were recruited by BearingPoint, Inc., under contract [8]. BearingPoint implemented the study in early December 2005 at the National Cancer Institute usability labs (Rockville, MD) and their home offices (McLean, VA). They also performed the preliminary data analysis. No more than nine participants took part in each unique set of pre-assigned tasks, thus US Office of Management and Budget clearance was not required, in compliance with regulations governing federal data collection under NIH policy 1825. The study fulfilled the conditions for IRB review exemption. All participants were compensated.

Recruitment was guided by three variables (Table 1):

- Language: Spanish, mono-/bilingual
- Age: younger (20-45 yrs.)/older (>45 yrs.)
- Health Literacy: lower/higher

Each part of the study involved at least one representative from each of eight subgroups (2 Language values x 2 Age values x 2 Health Literacy levels). Parts 1 and 3 involved 8 participants each and part 2, 20 participants, for a total of 36.

Variables	Values	No. (n=36)
Language	Monolingual	18
	Bilingual	18
Age (years)	20-45 yrs.	17
	>45 yrs.	19
Health Literacy	Lower	18
	Higher	18

Table 1. Participants by variables.

Candidates from Hispanic community centers and community-based organizations from Maryland, Virginia, and Washington, D.C. were screened to ensure a mix of participants by each variable. Health literacy level was determined by the Spanish vocabulary and language skills demonstrated during the screening interview, and by the number of questions answered correctly on the Short Assessment of Health Literacy in Spanish-speaking Adults (SAHLSA) [9], a new instrument validated on the Rapid Estimate of Adult Literacy in Medicine (REALM) [10].

Data Collection

- *Semi-Structured Interviews* (one-on-one) to assess users' needs and expectations (part 1), and elicit feedback for improvements (part 3).
- *Usability Testing* (part 2) to assess satisfaction with the Spanish-language prototype, ease of navigation, and users' comprehension.

Part 1: The interviews (n=8) sought to assess users' needs and expectations when seeking information about clinical trials, generally. The interview probed participants' awareness of clinical research concepts, including depth of understanding and past experience. After being given a hypothetical situation where clinical trial information might be needed, participants were handed a relevant printed doclet for comment (e.g., "Is this what you would expect?" "What other information would you like to see?").

Part 2: Participants (n=20) were asked to complete different unique scenario-based tasks online while describing their actions ("think-aloud" protocol [11]). For example, one set of tasks asked one group of users to find an asthma inhaler trial and determine its purpose, starting date of the trial, and eligibility requirements. The scenarios were pilot tested for clarity and internal validity. The order of the scenarios was randomized among users to control for order bias. Access to the Spanish prototype was provided through a password-protected URL. While one researcher facilitated the session using a prewritten script, another observed the information seeking behavior (e.g., navigation paths, potential confusion or problem areas, recovery). Sessions ended with an open interview to obtain overall impressions of the prototype, feedback, and whether they would revisit the site.

Part 3. Semi-structured interviews (n=8) were conducted to assess satisfaction with the doclets by themselves, and compared to doclets+purpose and to the full-text English-language documents. A scenario set the context for reviewing a printed doclet. After doing so, participants were asked to describe questions they might have, comment on particular helpful information, satisfaction with the amount of information provided, and perceived ease of understanding. They were then shown the equivalent information in the two other formats and asked how well each of these met their needs.

RESULTS

Both qualitative and quantitative methods were used to evaluate aspects of information seeking needs for clinical research studies and the Spanish ClinicalTrials.gov prototype.

Part 1. Comprehension and Expectations

Overall, while most participants had searched the Internet for health information in Spanish recently, they were only somewhat familiar with clinical research concepts. Participants frequently cited several data items as being useful in a clinical trials health information system (Table 2).

Rank	Clinical Trials Information Needs	n=
1	Benefits and risks of participating	7
2	Purpose of the study in simple terms	5
3	Treatment type (medicines, placebo, etc.)	5
4	Sponsoring organization (for reliability)	5
5	Duration of study participation	5
6	Location where the study is conducted	3
7	Assistance offered during participation (i.e. transportation, living arrangements)	2
8	Probabilities of improving a medical condition as a result of trial participation	2

Table 2. Top eight items considered useful, rank ordered by number of respondents.

Nearly all participants commented on the need for some information about the purpose in the doclet. Other study details, such as study design and type, were not deemed helpful:

> The study type of this trial is observational, but aren't all studies observational?

Language	Monolingual				Bilingual			
Health Literacy	Higher		Lower		Higher		Lower	
Age	Younger	Older	Younger	Older	Younger	Older	Younger	Older
Total Participants: 20	n=2	n=2	n=3	n=3	n=3	n=3	n=2	n=2
All Tasks: Fully Completed	75.0%	25.0%	25.0%	33.4%	91.7%	87.6%	71.4%	75.0%
Partially Completed	0%	25.0%	41.7%	33.3%	8.3%	0%	26.6%	0%

Table 3. Task completion rates during the usability testing (part 2) by demographic variable.

Part 2. Usability Testing

Participants generally found the prototype easy to navigate and search, and approved of the content:

> A very interesting and informative Website, with different ways of searching for specific information
>
> I liked the information in the links at the top.
>
> What I liked the most is that it is in Spanish and I can understand it.

While several commented on the high quality of the translation, they also stated that would prefer less technical descriptions. In general, older monolingual users with lower health literacy had lower task completion rates (Table 3) due to greater difficulty navigating, finding, and understanding information. This may result from less online information seeking experience than participants in other groups. Thus, greater attention is needed in (1) providing domain-specific terminology appropriate for consumers and (2) assisting older monolingual users, inexperienced in online health information seeking. Specific usability problems at the prototype Website include:

- Lack of spell checking (e.g., "tiroide" rather than "tiroides" returned no documents)
- Confusion about "advanced" features (e.g., Search Within Results, Query Details)
- Too many studies in the search results list
- Lack of US state names in List by Map

Part 3. Satisfaction with Doclets

Overall, while participants agreed that clinical trials information in Spanish is important; they felt that the doclet design did not provide enough information to determine the purpose of a study. Others mistakenly thought that the heading, *Finalidad* ("Purpose"), was a link to a purpose statement in Spanish.

In comparing doclets with doclets+purpose one participant observed:

> …the doclets without the purpose section gives the impression that this study is like a program to lose weight whereas the one with the purpose section seems more to be about research taking place.

Several participants disliked the description of the study type (e.g. factorial assignment, natural history) without explanation or definitions, because:

> …including the study type without explaining the concepts creates confusion, as it raises more questions than it proposes to answer.

Finally, none of the participants used the Spanish-language link (in the Purpose section) to the corresponding full-text English record: they either did not realize that clicking it would provide the purpose, albeit in English, or did not want to be linked to an English-language webpage. Even bilingual users stated they would not want to switch languages while using a Spanish-language Website.

DISCUSSION

The results of this user study suggest that combining cross-language retrieval and display strategies may be a resource-effective way to make English-language consumer health information systems more accessible to users whose dominant language is not English. A recent survey indicated that while 15% of respondents have participated in clinical studies, 63% said they would be "likely" to participate [12]. Based on our results, access to study information may be a major barrier for monolingual and bilingual Spanish-speakers. To be effective, the sites need to define concepts in lay terms, provide sufficient context, and be simple to navigate and use.

While participants across the spectrum appreciated the Spanish prototype as a "first step," they indicated a strong preference for some information about the purpose of the trials in the doclets. That the Spanish-language link to the English-language record was not used or felt to be useful was an unexpected finding. To encourage its use, the link label might be changed to suggest (in Spanish) that a copy of the full English document be printed out, to provide healthcare providers with more information about the trial. Other places on the Website might also remind users that only partial information is available in Spanish and decisions about trials should be made with the full information, in consultation with professionals. In general, the consensus among participants regarding information display was:

- Doclets do not convey the gist;
- Doclet+Purpose is better than the doclet alone;
- Lengthy/technical descriptions in the full-text English records are not helpful.

The doclet+purpose alternative appears to be a reasonable compromise under the assumption that more contextual information is better for the end

user, rendering the Spanish display closer to the full-text English documents, which were our reference standard both for retrieval results and display. The doclet alone does not provide end-users with enough information for their needs, and the free-text Purpose field in the English record is not amenable to controlled-vocabulary look-up. Translating an excerpt from the purpose description may provide sufficient information about the study without consuming substantially more resources, especially with the text extraction algorithm. However, the doclet+purpose design needs further study to assess whether this format is indeed an optimal presentation for helping consumers find data about clinical trials.

Future research includes exploring the generalizability of this approach to other consumer health systems (e.g., Genetics Home Reference) and languages (e.g., French). We will also continue to work on natural language processing and informatics tools to populate/curate the bilingual tables in the CLIR module used to convert query terms from Spanish into English. We also plan to investigate ways to incorporate a spell-checking mechanism for Spanish queries.

Finally, from a development perspective, a number of issues remain regarding converting the Spanish ClincalTrials.gov prototype into a production system Current challenges include the need for additional technical resources, Spanish-speaking user support personnel and materials, ongoing maintenance of the translation tables, extraction and translation of purpose statements for all doclets, and continued synchronization of new and modified clinical trials.

CONCLUSION

Our preliminary data suggest that combining CLIR and target language display techniques increases accessibility to existing English-language online consumer health information among non-English speaking users. Nevertheless, such an approach is only a partial solution to the general problem of multilingual information access. Monolingual or multilingual systems that operate entirely within their own languages are likely to provide optimal performance. However, the cross-language model may temporarily bridge language gaps between resources and users until online health resources in other languages are more widely available.

ACKNOWLEDGMENTS

We thank Gabrielena Alcala Levy and the team from BearingPoint, Inc for implementing the study under an NIH Office of Evaluation Set-Aside Funding Award (05-126-NLM), Daniel Lee Shoou-Yih et al. (University of North Carolina at Chapel Hill) for support with SAHLSA, and Laurel Graham for valuable help. This work was supported by the Intramural Research Program of the NIH, NLM.

REFERENCES

[1] Hesse BW, Nelson DE, Kreps GL, Croyle RT, Arora NK, Rimer BK, et al. Trust and sources of health information: the impact of the Internet and its implications for health care providers: findings from the first Health Information National Trends Survey. Arch Intern Med. 2005 Dec 12-26;165(22):2618-24.

[2] Pew Research Center. Trends 2005: Information for the Public Interest. Washington, DC: Pew Research Center; 2005.

[3] Pew Internet and American Life Project. “Hispanics and the Internet,” July 2001.

[4] Halsted DD, Varman B, Sullivan M, Nguyen L. Consumer health information for Asians (CHIA): a collaborative project. J Med Libr Assoc. 2002 Oct; 90(4):400-5.

[5] Rosemblat G, Graham L. Cross-language search in a monolingual health information system: flexible designs and lexical processes. Proc Ninth ISKO Conference 2006: 173-182.

[6] Rosemblat G, Tse T, Gemoets D, Gillen JE, Ide NC. Supporting access to consumer health information across languages. ICML9. Brazil. 2005 Sept. http://www.icml9.org/program/track6/

[7] Morris AH, Kasper GM, Adams DA. The Effects and Limitations of Automated Text Condensing on Reading Comprehension Performance. Information Systems Research. 1992; 26: 17-35.

[8] BearingPoint, Inc. Focused process evaluation to assess the usefulness and effectiveness of experimental displays of information about clinical trials for Spanish-speaking consumers: final report. Submitted to LHNCBC, NLM, NIH. McLean, VA: BearingPoint, Inc.; 2006 Feb.

[9] Lee S-YD, Bender DE, Ruiz RE, Cho YI. Development of an easy-to-use Spanish health literacy test. Health Services Research. In press.

[10] Davis TC, Crouch MA, Long SW, Jackson RH, Bates P, George RB, et al. Rapid assessment of literacy levels of adult primary care patients. 1991. Fam Med. 23(6):433-5.

[11] Lewis C, Rieman J. Task-Centered User Interface Design: A Practical Evaluation. 1993. [Accessed at http://hcibib.org/tcuid/ on June 22, 2006]

[12] Charlton Research Company for Research! America. National Survey. 2006. [Accessed at http://www.researchamerica.org/polldata/2006/NationalPoll2006.pdf on June 22, 2006]