The Video-Based Communication Assessment: An Innovative System for Assessing Clinician-Patient Communication

Kathleen M. Mazor, Ann M. King, Ruth B. Hoppe, Annie O. Kochersberger, Jie Yan, Jesse D. Reim
Background. Good clinician-patient communication is essential to quality health care, and key to patient centered care. However, individuals and organizations seeking to improve in this area face significant challenges. A major barrier is the absence of an efficient system for assessing clinicians’ communication skills and providing meaningful, individual-level feedback.

Objective. Our objective was to design and create the Video-Based Communication Assessment (VCA™), an innovative, flexible online system for assessing and ultimately enhancing clinicians’ communication skills.

Methods. We began by elaborating a vision of the VCA. Specifically, we determined that the assessment system should be: 1) convenient and efficient, accessible via computer, tablet or smartphone; 2) case-based, using video patient vignettes to which users respond as if speaking to the patient in the vignette; 3) flexible, allowing content to be tailored to the purpose of the assessment; 4) incorporate the patient’s voice by crowdsourcing ratings from analog patients; 5) provide robust feedback including ratings, links to highly rated responses (i.e., exemplars), and learning points; and 6) ultimately have strong psychometric properties. We collected feedback on the concept and then proceeded to develop the system. We identified several important research questions which will be answered in subsequent studies.

Conclusions. The VCA is a flexible, innovative system for assessing clinician-patient communication. It enables efficient sampling of clinicians’ communication skills, supports crowdsourced ratings of these spoken samples, and multi-faceted feedback reports.
Introduction

Good clinician-patient communication is essential to quality health care and a key element of patient-centered care.[1, 2] There is a substantial and growing evidence base documenting the critical importance of effective clinician-patient communication for a variety of health outcomes.[3, 4] Training in how to communicate effectively with patients and families is required for medical school and residency program accreditation,[5, 6] and competency in communication is a requirement for licensure.[7, 8] Financial incentives for excellent clinician-patient communication and penalties for poor communication are becoming widespread.[9-11] Many practicing physicians are finding that a portion of their compensation is dependent on their patients’ perceptions of their communication skills, and healthcare reimbursement rates are increasingly influenced by patients’ ratings of communication.[12-14] In short, research evidence in addition to societal, policy, and financial forces have converged to influence healthcare systems and individual clinicians to value communication, and to seek to improve in this area.

While the importance of high-quality clinician-patient communication is widely acknowledged, individuals and organizations seeking to improve in this area face significant challenges. One barrier to improvement is the lack of assessment tools which can produce timely, specific, and individual-level feedback. Patient satisfaction and patient experience surveys are widely used in practice, but often entail a significant delay between the time of assessment and provision of feedback. In addition, such measures often have relatively few items assessing communication, and so do not provide the degree of specificity needed to support improvement. Most medical schools and some residency programs utilize standardized patients and simulated encounters for both formative and summative assessments, but such simulations are expensive and time consuming. Thus, there is a need for an efficient approach to assessing clinicians’ communication skills.

Recognizing the lack of an efficient assessment in spite of the critical importance of communication in medicine, the National Board of Medical Examiners (NBME) invested in the development of a communication skills assessment system. This paper describes the development of the Video-Based Communication Assessment (VCA™), an innovative, flexible online system for assessing and ultimately enhancing clinicians’ communication skills.

Methods

Overview. The VCA is a flexible online system for collecting samples of how clinicians communicate, obtaining ratings of these communication samples, and summarizing the ratings as feedback. Users can access the assessment online or through the VCA app. Brief (10-20 second) patient vignettes are presented to users, who are instructed to respond as if they were speaking to the patient. Vignette content is flexible, and can be crafted to focus on general communication skills (e.g., information gathering) or specific communication situations (e.g., delivering bad news or discussing breast cancer screening). Users’ spoken responses are captured and presented to online raters using crowdsourcing. How responses are rated is also flexible, and can target specific skills (e.g., information clarity). When ratings are complete, feedback reports are generated, displaying the user results in the context of his or her cohort, highly rated responses (i.e., exemplars), and learning points detailing rater feedback. The flexibility inherent in the VCA means that assessments can be created that apply to a variety of clinician types (e.g., physicians, pharmacists, nurses, health coaches), specialties (e.g., primary care, oncology, surgery, palliative care) and training level (e.g., medical students, residents, practitioners).

In the sections below we describe the VCA, and each of its components.

VCA Vision. Because the overarching goal of this project was to develop a formative tool that would be useful to busy clinicians, we sought to make the VCA short and efficient. To make it convenient for clinicians, it would need to be easy to access and available whenever the clinician has time. To be engaging and somewhat realistic, it should be case-based. To be broadly useful, it should be flexible with regard to content in order to allow assessment of a broad range of communication skills. It would also
need to include a strong feedback component. Consistent with NBME standards and practice, we sought to create a system that would result in assessments with strong psychometric properties. In recognition of the importance of the patient’s voice in assessing patient-centered communication, we determined that analog patients (i.e., lay people who take on the role of the patient in a simulation) would be included in the assessment process. Finally, we sought to design an assessment that would be affordable not just for educational programs, training programs, healthcare organizations and systems, but also for individual clinicians. We began the instrument development process by specifying how each of these properties could be operationalized in the VCA. (Table 1)

Table 1. VCA Assessment Properties and Operationalization

<table>
<thead>
<tr>
<th>Convenient and efficient</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Accessible 24/7</td>
<td></td>
</tr>
<tr>
<td>Allow participants to pause and return if interrupted</td>
<td></td>
</tr>
<tr>
<td>Easy to use interface with intuitive navigation</td>
<td></td>
</tr>
<tr>
<td>Accessible online or through the VCA app</td>
<td></td>
</tr>
<tr>
<td>Cases would incorporate “slices” of patient encounters rather than entire encounters</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case-based</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Present brief patient vignettes of realistic communication situations</td>
<td></td>
</tr>
<tr>
<td>Capture participant’s spoken response</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Flexible content</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>The VCA approach will provide a methodology and a template for communication assessments, which will be populated with a variety of clinical vignettes in order to assess different communication skills, content, or specialty-specific skills.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Incorporate the patient’s voice</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Analog patients are engaged to rate participants’ spoken responses</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Robust feedback</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>The VCA will include a robust feedback component, collecting ratings and generating feedback within a relatively short timeline. Ratings will be provided for each vignette for individual users and for cohorts of users (enabling peer comparisons and benchmarking). Feedback reports will allow users to review the vignette materials and their own responses, and will include audio-recordings of exemplary responses (i.e., those responses to each vignette which were among the most highly rated). Curated “learning points” culled from analog patients’ comments about what constitutes an optimal response to a given vignette will be included.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strong psychometric properties</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Refer to a six function model of communication as the underlying framework[8]</td>
<td></td>
</tr>
<tr>
<td>Engage practicing providers in vignette development.</td>
<td></td>
</tr>
<tr>
<td>Use multiple vignettes in each assessment, and multiple raters of each response to increase score reliability. Conduct reliability and validity studies.</td>
<td></td>
</tr>
</tbody>
</table>

**Concept Development**

We envisaged the VCA as an online assessment tool which registered users access by logging on to a dedicated website via laptop or desktop computer, or through an app on smartphone or tablet. After logging on, users would be presented with brief clinical background information (in text) which provides the context for a clinical encounter. Users would then click to play a brief video of the patient in the encounter. The video would end at a point where the provider would be expected to speak to the patient, and the user is prompted to respond (i.e., “What would you say next?”). The user would respond, speaking as if he or she were actually speaking to the patient. This spoken response would be audio-recorded and stored. The sequence of reading brief text providing context, watching a brief video, and responding as if one were talking to the patient in the encounter would be repeated for multiple vignettes.
When a cohort of users had completed the assessment, responses would be rated by analog patients using crowdsourcing. When ratings were completed, feedback reports would be created. These would include individual user’s ratings, comparative data on the user’s cohort, learning points derived from analysis of crowdsource raters’ comments on what would constitute a satisfactory response to the patient, and exemplary, highly rated responses which the user could compare to his or her own response.

Figure 1. Overview of the VCA Process

**Concept Testing.** We created a brief presentation that described the VCA, and used this as the basis for soliciting feedback from providers, educators and health system leaders in order to make an early determination of the potential attractiveness and usefulness of the VCA. Reactions were strongly positive.

We created sample vignettes, writing brief text to provide the clinical context, and producing amateur videos. We integrated these two components and the stimulus question (“What would you say next?”) into PowerPoint, and using separate recording devices, collected spoken responses from a small convenience sample of clinicians. We then piloted the use of crowdsourcing to gather ratings of these spoken responses in an online survey administered via Amazon Mechanical Turk (MTurk). Results were positive; we were able to elicit spoken responses, and to obtain ratings within hours of posting on MTurk. We therefore proceeded with development.

**Results**

**VCA Content.** The VCA is not a single, immutable assessment, but rather an assessment system that can be adapted and tailored to assess a variety of communication skills. Much like standardized patient-based assessments where case content is varied depending on the purpose of the assessment, VCA vignettes can be created to assess specific skills. For example, one set of VCA vignettes could be created to assess skill in providing information in a general medicine outpatient context, while another set of
vignettes could be created to assess skill in disclosing medical errors. Our goal in creating the first set of VCA vignettes was to assess communication skills broadly, using clinical situations that would be familiar and relevant to providers from a variety of backgrounds.

**Vignette Development.** We engaged a multi-disciplinary panel of clinicians and educators to participate in a one-day vignette development workshop. After the panel, the authors used an iterative process to refine the vignettes developed or suggested during the panel, and to generate additional vignettes. The resulting vignettes were reviewed by the authors, who assessed the video portrayals for realism and whether the produced vignette would be likely to appropriately stimulate a spoken response that could in turn be rated by analog patients. An example of a vignette (with a still photo representing the video) is presented in Figure 2.  

**Figure 2. Screen shots of VCA user interface via computer and app**

**Rating Process.** Consistent with our intent to incorporate the patient voice into the assessment, the VCA is designed to allow crowd-sourced analog patients to rate the users’ spoken responses. Analog patients are lay people who take on the patient’s perspective and rate the encounter as if they were the actual patient. Analog patients are increasingly used to study and assess clinician-patient communication, and a growing body of evidence suggests that analog patients’ ratings are reliable and valid. Indeed, analog
patients’ ratings may be more informative than actual patients’ ratings, as they avoid the ceiling effects that are often observed when the latter are used. [15]

Analog patients can be recruited in a variety of ways, from a variety of sources. Initially, we recruited analog patients using crowdsourcing via Amazon Mechanical Turk (MTurk). MTurk is a widely-used online workplace that enables requesters to utilize crowdsourcing to complete specific tasks. Requesters can constrain which workers may complete a task by specifying eligibility criteria (for instance, to workers who have consistently demonstrated a high degree of accuracy on prior tasks), or by assessing qualifications through screening questions (for instance, a requester might accept those who report they have had a doctor’s appointment for a specific condition in the prior year). There is a growing body of literature describing the demographic characteristics of MTurk workers, the quality of data collected via MTurk, and methods for improving data quality. [16]

At present, analog patients recruited for the VCA are first oriented to the task, and instructed to imagine themselves as the patient in the situation to be presented. They then view brief text describing the clinical context for one vignette in lay language, followed by a video of the patient speaking to the provider. He or she then listens to a recorded response, rates the response on 2-6 items, then proceeds to the next recorded response and repeats the process. The number of responses that will be bundled together (i.e., a Human Intelligence Task or HIT as termed in Amazon Mechanical Turk) is expected to vary between 10 and 20, with the optimal number or range to be determined empirically as data are collected. Each provider’s spoken response will be rated by multiple analog patients; our prior work suggests that approximately 15 analog patients will provide sufficient reliability, but the number of raters will also be determined empirically, and with consideration of the purpose of the assessment. The number of analog patients needed may also vary depending on the specific vignettes used, the providers involved, and other factors; these influences will be explored in future research.

After rating all audio recordings in the bundle, analog patients are asked to respond to a single open-ended item “What would you have wanted your provider to say if you had been in this situation?” Responses to this open-ended question will be reviewed by VCA staff to identify prominent themes. For each vignette, themes will be translated into key learning points for inclusion in feedback reports.

**Rating Items.** Because the VCA process is very different from typical communication assessment processes, the items which analog patients use to rate users’ responses are of critical importance. We referred to three sources in developing draft rating items: 1) the six-function model which provides the framework for communication assessment used by the NBME; 2) the rating items which were used in an earlier communication assessment which closely resembles the current VCA[17]; and 3) the Consumer Assessment of Healthcare Providers and Systems (CAHPS) item sets, which are an increasingly important point of reference for many providers and health care organizations.[13] A set of six items was created for pilot testing. Items will undergo extensive testing and psychometric analyses. An example of the type of item being tested is “The provider explained things in a way I could understand.”

**Flexibility with respect to raters and ratings.** While we anticipate that analog patients using rating items developed and tested for the VCA will be the primary way that responses are scored, the VCA system is designed to be flexible and to accommodate other raters and rating items. For example, researchers have expressed an interest in using the VCA to efficiently collect samples of clinicians’ typical ways of communicating with specific types of patients. Such responses would then be accessed and scored by the research team according to criteria specific to the research question. In this case raters could be analog patients, trained research staff, or experts selected for their expertise in a particular content area.

**Feedback**

Because we anticipate that a major application of the VCA will be as a formative assessment, feedback is of fundamental importance. Two types of feedback reports are possible: 1) individual feedback reports
which will provide detailed results for a single user, with aggregate, de-identified results of his or her cohort provided for comparison; and 2) organization-level feedback reports which will provide detailed and summary results for an entire cohort of users.

Individual feedback reports will contain both quantitative and qualitative components, as summarized in Table 2.

**Table 2. Elements of Individual Level Feedback Reports**

| Overall rating | Individual’s overall mean rating, aggregated across all items, raters, and vignettes  
Cohort overall mean rating, aggregated across individuals, items, raters and vignettes |
|----------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Vignette ratings | Individual’s vignette level mean rating, aggregated across items and raters for each vignette.  
Cohort vignette level mean rating, aggregated across individuals, items and raters for each vignette. |
| Vignette learning points | Curated learning points for each vignette, extracted from content analysis of analog patients’ comments, with illustrative quotes |
| Exemplary audio-recordings | For each vignette, one or more highly rated audio responses will be posted to illustrate the learning points, and to provide a realistic example of how a provider might respond in the situation portrayed in the vignette. |
| Individual responses and vignette materials | Each individual will be able to review the vignettes (i.e., clinical context and patient video) and their own responses (i.e., audio recordings) within the feedback report. |

An individual user will be able to share links to his or her personal feedback report at his/her discretion. For instance, a user could choose to share a link with a local coach who would work directly with him or her in one-on-one sessions discussing the feedback reports, and reviewing his or her responses and comparing these to the exemplary responses and comments.

Organizational-level feedback reports will include all of the components of the individual feedback reports, including ratings, spoken responses, exemplars, and analog patient comments. In addition, these reports will allow supervisors to review the relative standing of individuals in a cohort. In the future, as the database of responses grows, multiple benchmarks will be available, showing, for instance, the average overall performance of Internal Medicine residents or Family Medicine physicians who have completed the VCA. At present, organizational level feedback reports do not contain the names of the individual users; how this information could be included in future versions with consent of users.

**Future Research**

We have identified a number of research questions related to both operational and psychometric considerations of the VCA (see Table 3). Data collection related to research questions 1-4 is currently underway. The answers to these questions will inform the final design and implementation. However, because the VCA is a system rather than a single assessment form, most of the research questions below will have conditional rather than absolute answers.

**Table 3. Priority Research Questions and Corresponding Research Strategies**
Research Question 1: Does the VCA (including the user interface, assessment process, and feedback reports) meet the needs of users and customers?

Research Strategy: Brief post-assessment surveys; user and customer interviews, market research?

Research Question 2: How many vignette responses, rating items, and raters will be needed to obtain a generalizability coefficient (g) of .80 or higher?

Research Strategy: Generalizability studies

Research Question 3: How does the wording of the items presented to analog patients affect ratings, and what items result in psychometrically sound scores?

Research Strategy: Sequential testing of various items and response options with independent samples of analog patients rating the same responses

Research Question 4: To what extent are analog patient characteristics (e.g., age, gender, race/ethnicity, education, geographic residence) associated with differences in ratings?

Research Strategy: Statistical analysis of the impact of specific analog patient characteristics on ratings, and assessment of the interaction between analog patient characteristics and vignette characteristics.

Research Question 5: Are scores on the VCA correlated with other measures of clinician-patient communication, patient experience, or patient satisfaction?

Research Strategy: Correlational studies comparing users’ scores on the VCA with scores on relevant items from measures collected in routine practice (e.g., CAHPS scores)

Research Question 6: Does participating in VCA contribute to improved provider performance?

Research Strategy: Pre/post studies of VCA users scores on measures collected in routine practice.

Discussion

The VCA is a new system for assessing clinician-patient communication with several innovative features. First, it utilizes a “thin-slice” approach to assessing communication skills which allows efficient sampling of a variety of encounters and moments within encounters, thereby providing a rich and robust stimulus. Second, it is accessible to clinicians on their digital devices at any time, in any location with internet connectivity. Third, assessments are designed to be brief. These features address a major impediment to clinician assessment and training: the lack of time and the need for clinicians to go to a specific location (e.g., simulation center) for assessment—especially for those who are post-licensure or certification. A third innovative feature of the VCA is our focus on spoken communication. To our knowledge, the VCA is the first digitally-based assessment of clinician-patient communication which captures the clinician’s spoken response. Fourth, the VCA allows incorporation of the patient’s voice into the assessment process. While patient experience surveys also incorporate the patient’s voice, the VCA approach, where the specific communication situations (i.e., the patient vignettes) are standardized within an assessment form, provides a stronger basis for comparisons, and allows clinicians to review the specific situation and response. Finally, feedback can be provided relatively more quickly than is typical of patient surveys.

The VCA has several other attractive features. Both the vignette content and rating approach are flexible and easily modified to meet different assessment needs. Content can be created to target different types of clinicians (primary care, surgical, subspecialty), those in training, and other types of health care workers, e.g., nursing or pharmacy. In addition, assessments of inter- or intra-team communication could easily be developed using the VCA. With regard to rating flexibility, while our focus to date has been to gather
online ratings using crowdsourcing, other raters, such as supervisors, content experts, or trained research staff could be used to generate ratings and feedback.

The VCA is an assessment tool which could be useful to individuals and organizations seeking to improve clinicians’ communication skills, evaluate the effectiveness of training programs, or document proficiency in communication. For instance, a hospital’s chief of medicine might use VCA results to identify those clinicians most in need of remediation. VCA results could be used to identify the most skilled clinicians, in order to engage those clinicians as peer coaches. VCA results could also be used to identify particular areas of strength or weaknesses for a given clinician—e.g., a clinician might be identified as needing additional support in how to disclose medical errors. Initial feedback on the VCA suggests that such inferences and decisions are relevant and important to hospital leaders, residency directors, and others in similar positions.

The VCA has important implications for the research setting, as it enables efficient, remote, targeted assessment of clinicians’ communication skills in the context of descriptive and intervention studies. Vignette-based surveys have been widely used in research to explore how clinicians respond to variations in clinical situations and patient characteristics, and how patients respond to variations in clinicians’ communication strategies. Most often, the vignettes presented are text-based, and the responses are collected via rating scales. While some studies, particularly those using analog patients, have used video vignettes as stimuli, we are not aware of any studies (with the exception of our own prior work) that have captured spoken responses. There are important differences in spoken and written communication.

The VCA will allow researchers to efficiently study not only what clinicians say, but also how they say it, and how patients respond to these communicative acts. Rigorous research in this area could ultimately provide new insights into what constitutes good communication from the patient’s point of view, as well as a more complete understanding of the extent to which patients’ perceptions of communication vary according to their personal characteristics.

There are a number of unanswered questions about the VCA and its impact. Specifically, further work is needed to determine whether use of the VCA, and the feedback reports in particular, result in behavior change, and whether any positive impact is obtained through the use of the report alone, or whether the support of a coach or teacher is required. Another uncertainty relates to the potential for widespread adoption of the VCA. While our preliminary conversations with potential customers and users were positive, it is not yet known whether institutions, organizations, programs and individuals will find this tool attractive or useful.

While the VCA has many attractive features, it also has limitations. The use of “slices” of communication behaviors allows efficient assessment, but the lack of sustained interactions might also be considered a limitation. Our decision to use this approach was in part pragmatic—doing so allows efficient assessment using a variety of vignettes in a relatively short amount of time. We recognize that this approach does not provide an assessment of clinicians’ ability to adjust their communication skills over the course of an interaction, or to develop a relationship with a patient over time. At the same time, prior research suggests that this approach will result in valid and reliable scores; and that even brief “slices” of communication behaviors are informative and predictive of important outcomes.

We plan to conduct a series of studies to investigate the properties of VCA scores.

In summary, the VCA is a flexible, innovative system for assessing clinicians’ communication skills. Research is currently underway to provide insights into the reliability of VCA scores under various conditions, and to examine the impact of rater characteristics. We believe that the properties of the VCA will enable this new system to make a substantial contribution to the assessment of communication skills and ultimately to improving clinician-patient communication.