Short communication

The use of video vignettes to measure health worker knowledge. Evidence from Burkina Faso

Sheheryar Banuri[a], Damien de Walque[b,*], Philip Keefer[c], Ousmane Diadie Haidara[d], Paul Jacob Robyn[d], Maurice Ye[e]

[a] University of East Anglia, UK
[b] Development Research Group, The World Bank, USA
[c] Institutions for Development Department, Inter-American Development Bank, USA
[d] Health Nutrition and Population, The World Bank, USA
[e] Centre de Recherche en Santé de Nouna, Burkina Faso

ARTICLE INFO

ABSTRACT

The quality of care is a crucial determinant of good health outcomes, but is difficult to measure. Survey vignettes are a standard approach to measuring medical knowledge among health care providers. Given that written vignettes or knowledge tests may be too removed from clinical practice, particularly where "learning by doing" may be an important form of training, we developed a new type of provider vignette. It uses videos presenting a patient visiting the clinic with maternal/early childhood symptoms. We tested these video vignettes with current and future (students) health professionals in Burkina Faso. Participants indicated that the cases used were interesting, understandable and common. Their performance was consistent with expectations. Participants with greater training (medical doctors vs. nurses and midwives) and experience (health professionals vs. students) performed better. The video vignettes can easily be embedded in computers, tablets and smart phones; they are a convenient tool to measure provider knowledge; and they are cost-effective instruction and testing tools.

1. Introduction

The quality of care is a crucial determinant of good health outcomes, but is difficult to measure (Hrisos et al., 2009). In particular, clinical quality is more difficult to measure compared to structural quality of the facilities. Facility surveys, knowledge tests and direct observations of medical consultations all exhibit limitations, ranging from bias to cost, that are particularly pronounced in poor country settings. To address these limitations, we introduce a new measurement device based on clinical video vignettes. We developed to assess the quality of care among medical practitioners in Burkina Faso. Our measure is low cost, more immersive and potentially less prone to bias when incentivized.

Standard facility surveys generally include indicators of structural quality focusing on the availability of equipment, drugs, and commodities, as well as on staffing (training, experience and absenteeism). Structural quality, however, while a prerequisite for the quality of health services, does not accurately measure the care actually delivered by providers.

Different methods have been used to assess the quality of interaction between provider and patient (Leonard and Maestad, 2016). Patient exit interviews seek to obtain the patients' feedback about their experience. They can reveal useful information about respect, courtesy, or waiting times, but because patients are rarely well informed about standard medical protocols, exit interviews are usually less reliable about the content of care. They might also be affected by recall bias, if they are conducted long after the consultation, or by response bias, for example if conducted in or close to the health facility and where patients might feel embarrassed to express their true opinions.

Written survey vignettes are another standard approach to measuring the quality of interaction between provider and patient. They present surveyed staff with a typical case with symptoms and ask them to provide a diagnosis and recommend a treatment course (Glassman et al., 2000; Peabody et al., 2000, 2004a; Peabody et al., 2004b; Das and Hammer, 2005; Veloski et al., 2005). However, survey vignettes measure knowledge and not actual practice. They are not able to capture effort and therefore miss the "know-do" gap. In addition, written vignettes or knowledge tests may be too removed from clinical practice, particularly in poor countries. First, "learning by doing" may be an important form of training, so that, second, health workers might be uncomfortable with the written form.

Direct observations of medical consultations allow measurement of actual health care practice, but they are likely to be influenced by Hawthorne effects as providers increase their level of effort when they realize that they are being observed (Leonard and Masatu, 2005, 2006). The standardized patient method can address these concerns: actor patients, trained to feign a specific illness and record the provider's action, present themselves unannounced at a health facility (Das et al., 2012). However, this method is by its nature limited to health conditions which can easily be feigned and do not require an invasive procedure to be diagnosed. As such, it is quite difficult to apply to maternal and child health. Direct clinical observations and standardized patients are also relatively costly, logistically (hiring and training standardized patients can be a long process that is also difficult to scale given inconsistencies in performance, fatigue, etc.) and in monetary terms (standardized patients need specialized training which tend to increase the costs in terms of ability and time).

Acted-out case study vignettes are lower cost than standardized patient method and are similarly advantageous when health workers come from a variety of educational backgrounds. Acted-out vignettes measure the competence of health workers in a real clinical setting better than written vignettes (Leonard and Maestad, 2016). However, they still present similar, if lower, logistical and monetary burdens, compared to standardized patients.

We therefore developed a new type of provider vignette using videos presenting a patient with maternal/early childhood symptoms visiting the clinic. Compared to vignettes acted-out live, the video vignettes do not allow an assessment of the doctor-patient interaction. However, video vignettes confront substantially lower logistical obstacles. They are particularly easy to implement in contexts where tablet or phone-based surveys are now the norm. Compared to acted-out vignettes, they offer an exceptional ability to standardize presentation for measurement and research. Compared to written vignettes, they offer an opportunity to present complex signs and symptoms that might be hard to describe in a written vignette, but easier to demonstrate with an actor.

The video vignettes were applied within the context of a "lab-in-the-field experiment" in Burkina Faso that tests effects of various incentive measures on provider performance within the context of a "laboratory" setting. The lab experiment is one study component of a larger impact evaluation on a Performance Based Financing (PBF) pilot in Burkina Faso.

This paper describes our instrument and presents evidence of its validity. By validity, we refer to how well the video vignettes measure health worker knowledge (Sullivan, 2011). Since our interest was precisely in investigating the quality of care related to medical conditions that are difficult for actors to simulate, we could not easily validate the measure by comparing results with the standardized patient method. Instead, we focus on two necessary conditions for validity. First, participants with more experience and training should perform better than those with less. Second, and as a corollary, the measure should reflect the "real world" of significant heterogeneity of practitioner knowledge. Hence, practitioners should exhibit considerable dispersion in their performance as evidenced by variance in their responses. These validity conditions, linking the results of the vignettes to established measures of knowledge and capacity (medical qualifications and experience) are recommended in the literature (Downing, 2003; Cook and Beckman, 2006).

2. Methods

2.1. Recruitment

We recruited two types of participants in the study. First, in February–March 2014, we recruited 1029 medical professionals who were attending training sessions for a new program of PBF in the health sector. The sessions were organized at the regional level in the cities of Gourcy, Kaya, Koudougou, Nouna and Ouahigouya in Burkina Faso. At the sessions, participants were invited to take part in a lab-in-the-field activity which included the video vignettes. Each health facility (including primary care facilities and district/regional hospitals) in the particular region were expected to send at least one representative from their facility to attend the training sessions. Most of the facilities provide primary health care services.

Second, in April–May 2014, we recruited 1113 future health professionals (nurses, midwives, and doctors). Nursing and midwifery students were in their last year, and came from a public (Ecole Nationale de Santé Publique) and a private (Ecole de Santé Privée Sainte Edwige) nursing and midwifery school. Medical students were in their 5th or 6th year, and came from a public university (Faculté de Médecine, Université de Ouagadougou) in Ouagadougou, the capital city.

2.2. Video vignettes

We designed video vignettes representing cases related to maternal and early childhood care. The cases followed symptom description and treatment protocols as defined by Burkina Faso's national therapeutic guidelines as well as WHO IMCI guidelines. The cases were developed and tested with nurses at the two nursing schools in Ouagadougou in January 2014. Cases with adequate variation in responses were kept while others were replaced with new cases.

For ability measurement, we used 4 cases: a pregnant woman experiencing pre-term labor, a 6-month-old child with pneumonia, a lactating woman with mastitis and a 3-year-old child with measles. In order to capture the variety of ways in which patients present themselves to practitioners, each case consisted of a video lasting either 60 or 100 s. In all cases, the same actress was used to play the role of the patient, and always in the same setting (an average clinic). Two of the videos lasted 60 s each, and represented "simple" cases in which the actress used normal language and had a standard appearance. These cases were the pre-term labor and the mastitis cases. The 100 s videos represented cases with more symptoms and, in addition, the actress used rambling language and appeared to be poorer and less educated. This more disadvantaged socio-economic status was conveyed through the patient's dress and the vocabulary and language she used to describe her illness history and symptoms. The two 100 s cases were the child with pneumonia and the child with measles.

Table 1 reports the average time taken and the average score obtained for each of the four cases. The order of the 4 cases was set and not randomized. Since learning about the tasks was likely, we therefore expect the first case to have taken longer and to have yielded a lower score; this was indeed the case. In addition, subjects averaged less time on case 3, with a 60 s video, than for cases 2 and 4, with 100 s videos. The table indicates that the score increased across the four cases. While this steady increase was potentially driven by learning effects, it is important to emphasize that the four cases differed with respect to diagnosis, length and complexity, and are therefore not easy to compare.

The videos were shot and edited by locally-renowned director Boubakar Diallo. Care was taken to select an actress of neutral appearance in terms of ethnicity. Subjects viewed the video and were asked to select the correct (1) diagnosis, (2) treatment, (3) follow-up schedule, and (4) alternate treatment in case the patient did not respond well to the initial treatment course. For each question, they selected their answer from a multiple-choice list (see example in Fig. 1). Each answer set had one correct response, two nearly correct responses, and two wholly incorrect responses, yielding additional variation. "Wholly incorrect" responses would be inconsistent with the symptoms and data provided in the case while "nearly correct" responses would be consistent with most (but not all) of the symptoms provided to the health professional. Correct responses were those that followed the national guidelines.