The Impact of Perceived Frailty on Surgeons' Estimates of Surgical Risk

Mark K. Ferguson, MD, Jeanne Farnan, MD, MHPE, Josh A. Hemmerich, PhD, Kris Slawinski, MA, Julissa Acevedo, and Stephen Small, MD

Departments of Surgery and Medicine, Pritzker School of Medicine, Center for Research Informatics, and Department of Anesthesiology and Critical Care, University of Chicago, Chicago, Illinois

Background. **Physicians are only moderately accurate in estimating surgical risk based on clinical vignettes. We assessed the impact of perceived frailty by measuring the influence of a short video of a standardized patient on surgical risk estimates.**

Methods. **Thoracic surgeons and cardiothoracic trainees estimated the risk of major complications for lobectomy based on clinical vignettes of varied risk categories (low, average, high). After each vignette, subjects viewed a randomly selected video of a standardized patient exhibiting either vigorous or frail behavior, then reestimated risk. Subjects were asked to rate 5 vignettes paired with 5 different standardized patients.**

Results. **Seventy-one physicians participated. Initial risk estimates varied according to the vignette risk category: low 15.2% ± 11.2% risk; average 23.7% ± 16.1%; high 37.3% ± 18.9% ($p < 0.001$ by analysis of variance). Concordant information in vignettes and videos moderately altered estimates (high risk vignette, frail video 10.6% ± 27.5% increase in estimate, $p = 0.006$; low risk vignette, vigorous video 14.5% ± 45.0% decrease, $p = 0.009$). Discordant findings influenced risk estimates more substantially (high risk vignette, vigorous video 21.2% ± 23.5% decrease in second risk estimate, $p < 0.001$; low risk vignette, frail video 151.9% ± 209.8% increase, $p < 0.001$).**

Conclusions. **Surgeons differentiated relative risk of lobectomy based on clinical vignettes. The effect of viewing videos was small when vignettes and videos were concordant; the effect was more substantial when vignettes and videos were discordant. The information will be helpful in training future surgeons in frailty recognition and risk estimation.**

(Ann Thorac Surg 2014;■:■–■)

Judgment about the appropriateness of surgery or its anticipated outcomes is influenced by a variety of factors, some of which are surgeon related, including training level, recent adverse patient outcomes, and prioritization of risk versus benefit [1–4]. Major complications after lung resection lead to increased length of hospitalization and increased costs [5, 6]. Accurate risk estimation improves patient selection for surgery, obtaining informed consent, interventions for risk reduction, and allocation of appropriate perioperative resources.

Using vignettes of actual patients, practicing surgeons are more accurate than trainees in predicting risk, although neither group exhibits high accuracy [7]. That caused us to ask, how much of estimated risk is based on the surgeon's visual assessment of the patient (the "eyeball factor")? Seeing a patient has an important effect on how physicians view a patient's level of illness and risk in a clinical setting [8], but this effect has not been quantified in a surgical setting.

In this study we measured the degree to which the eyeball factor influences surgical risk estimation. We assessed the impact of a video of a standardized patient portraying vigorous or frail physical behavior on estimates of risk. Differences between practicing surgeons and trainees were also assessed. We hypothesized that trainees and practicing surgeons will differ in their risk estimates, that physicians' risk estimates will change significantly after they view a video of a standardized patient, and that the impact of videos on risk estimates will differ between these groups.

Material and Methods

A group of physicians and other geriatric specialists (M.K.F., J.F., K.S., S.S.), using an iterative process, developed a set of physical characteristics based on Fried's phenotypic criteria [9] that could be portrayed in a short, silent video, including weight loss, gait speed, strength, and fatigue (Appendix). Five standardized patients portrayed differing degrees of these physical characteristics of frailty, ranging from vigorous to frail, in videos in a mock outpatient clinic setting. Five geriatrics specialists rated the level of frailty portrayed in each of 62 videos. Two videos were selected for each standardized patient that had the highest interrater agreement for representing "somewhat vigorous" and "somewhat frail" behavior. Hereafter, these are referred to as "vigorous" and "frail" videos.

Accepted for publication April 4, 2014.

Address correspondence to Dr Ferguson, Section of Cardiac and Thoracic Surgery, University of Chicago, 5841 S Maryland Ave, MC5040, Chicago, IL 60637; e-mail: mferguso@surgery.bsd.uchicago.edu.

Published by Elsevier Inc

0003-4975/$36.00
http://dx.doi.org/10.1016/j.athoracsur.2014.04.141

Five short clinical vignettes were developed based on patients who underwent lobectomy for non-small cell lung cancer at our institution during 2011 and 2012 and were aged 50 to 69 years. The Charlson score (for assessment of risk status related to comorbidities) and EVAD score (for physiologic risk status specifically related to lung resection) were calculated for each patient [10, 11]; the possible range of Charlson scores is 0 to 37, and the possible range of EVAD scores is 0 to 12. In our dataset of 89 patients, Charlson scores ranged from 2 to 7 (mean 3.0 ± 1.2; median 3), EVAD scores ranged from 0 to 10 (mean 4.8 ± 2.8; median 4), and the range of combined scores was 2 to 15 (mean 7.8 ± 3.3; median 7). Patients were selected for vignettes based on the mean combined score plus or minus one standard deviation. Two patient vignettes were classified as low risk with total score of 4 or 5, one was classified as average risk with a total score of 8, and two patient vignettes were classified as high risk with a total score of 11. These scores were not revealed to the study subjects.

Thoracic surgeons serving as teaching physicians in traditional 2-year or 3-year US-based cardiothoracic training programs (n = 161) and cardiothoracic trainees in those programs (n = 125) were solicited to participate in the study using e-mail addresses provided by the Thoracic Surgery Directors Association. Participants provided their age, sex, comfort level with performing a lobectomy, and the current year of training or number of years of experience since completing training.

Participants were asked to read a clinical vignette and score the patient's risk for major postoperative complications after lobectomy from 0% to 100% using an anchored Likert-like scale. They then rated the importance of each of 13 clinical variables in their risk assessment for that patient. Participants then viewed one of two videos of the standardized patient paired with the clinical vignette, randomly selected between vigorous and frail behavior. Participants provided a second estimate of surgical risk based on the vignette and the video, rated the importance of the video in making the second risk estimate, and rated the importance of the physical behaviors portrayed in the video to their second risk estimate. This process was repeated for a total of five vignettes.

Study data were collected and managed using the REDCap (Research Electronic Data Capture) tool hosted at the University of Chicago. The REDCap tool is a secure, Web-based application designed to support data capture for research studies [12]. Independent and paired *t* tests, analysis of variance, and χ^2 analyses were used as appropriate. Analyses were performed using Minitab, version 16 (Minitab, Inc, State College, PA). The study was approved by our Institutional Review Board, and the requirement for written consent was waived.

Results

Subjects

Fifty-three thoracic surgeons and 18 cardiothoracic trainees participated in the study. Practicing surgeons were significantly older than trainees. There was no difference in sex distribution between the groups; both were predominantly male. Practicing surgeons had considerable clinical experience, and that likely resulted in the high frequency with which they rated their ability to perform a lobectomy as expert (Table 1). In contrast, none of the trainees rated their ability as expert ($p > 0.001$). The 71 participants completed scoring for 294 vignette-video pairings; 53 (75%) completed all five vignette-video pairings. The distribution of frail and vigorous videos was similar between the trainees and surgeons ($p = 0.411$). All scores were used in the analyses.

Initial Risk Estimates Based Only on Clinical Vignettes

Initial risk estimates, which were made without the influence of a paired video, are described in Table 2. When initial risk estimates for the five vignettes were combined into three risk categories, there was a significant difference among the risk scores. That illustrates good calibration of the participants' estimates relative to preassigned vignette risk category (Fig 1).

Some clinical factors in the vignettes (performance status, spirometry, diffusing capacity, smoking status, and possibly clinical impression) were instrumental in risk estimation, as determined by more than half of the sample rating them as important, whereas the majority of the elements were not considered important for most of the risk estimates (Table 3). The importance of clinical factors differed somewhat according to the level of risk evident in the vignettes. The impact of smoking status, obesity, and possibly cardiac status and diffusing capacity on risk estimates appeared to increase as vignette risk increased (Table 4).

Second Risk Estimates Based on Clinical Vignettes and Videos

Overall, second risk estimates made after viewing videos of standardized patients were similar to initial estimates (25.1 ± 16.2 versus 24.9 ± 18.1; $p = 0.863$). Initial and second risk estimates differed when assessed by vignette risk level. There was a significant increase in estimated

Table 1. Characteristics of Participating Surgeons and Cardiothoracic Trainees

Characteristic	Trainees (n = 18)	Practicing Surgeons (n = 53)	*p* Value
Age	35.5 ± 3.4	48.7 ± 9.6	<0.001
Male	11 (85%)	42 (72%)	0.361
Years cardiothoracic training or in practice	2.6 ± 1.2	13.3 ± 10.6	NA
Comfort level for performing lobectomy			<0.001
Naïve	0 (0%)	0 (0%)	
Learner	4 (22%)	1 (2%)	
Competent	14 (78%)	10 (19%)	
Expert	0 (0%)	42 (79%)	

NA = not available.