Society of Asian Academic Surgeons

Computed tomography in ventral hernia diagnosis: have we improved? A quality improvement initiative

CrossMark

Julie L. Holihan, MD, MS,[a,*] Deepa Cherla, MD,[a] Katherine J. Blair, MD,[b]
Steven S. Chua, MD,[b] Joseph P. Hasapes, MD,[b] Eduardo J. Matta, MD,[b]
Kaustubh G. Shiralkar, MD,[b] Venkateswar R. Surabhi, MD,[b]
Varaha S. Tammisetti, MD,[b] Tien C. Ko, MD,[a] and Mike K. Liang, MD[a]

[a] Department of Surgery, University of Texas Health Science Center at Houston, Houston, Texas
[b] Department of Radiology, University of Texas Health Science Center at Houston, Houston, Texas

ARTICLE INFO

ABSTRACT

Background: Previous studies suggest that agreement between readers of computed tomography (CT) scans for the diagnosis of a ventral hernia (VH) is poor (32% agreement, $\kappa = 0.21$). Recommendations were developed by surgeons and radiologists after determining common reasons for disagreement among CT reviewers; however, the long-term effect of adoption of these recommendations has not been assessed. The aim of this quality improvement (QI) project was to determine whether the incorporation of recommendations developed by surgeons and radiologists improves agreement among reviewers of CT scans in diagnosing a VH.
Methods: A prospective cohort of patients, with a CT scan of the abdomen and pelvis in the past 1 y, attending a surgery clinic at a single institution was enrolled. Enrolled subjects underwent a standardized physical examination by a trained hernia surgeon to determine the likelihood of a clinical VH (no, indeterminate, or yes). The QI intervention was the distribution and implementation of previously described recommendations. After a year of intervention, independent radiologists assessed patients' CT scans for the presence or absence of a VH. Percent agreement and kappa were calculated to determine interobserver reliability. In-person discussion on scans with disagreement was held, and the results were used as a "gold standard" to calculate sensitivity, specificity, positive, and negative predictive values for CT scan diagnosis of a VH.
Results: A total of 79 patients were included in the study. After QI intervention, seven radiologists agreed on 43% of the scans, and κ was 0.50 ($P < 0.001$). Agreement was highest among patients with a high clinical likelihood of a VH and lowest among patients with an indeterminate clinical likelihood. Sensitivity and specificity were 0.369 and 0.833, respectively.
Conclusions: After the implementation of recommendations, there is improved agreement among radiologists reading CT scans for the diagnosis of a VH. However, there is substantial room for improvement, and CT scans for the diagnosis of VH is not ready for widespread use.
© 2017 Elsevier Inc. All rights reserved.

* *Corresponding author.* Department of Surgery, University of Texas Health Science Center at Houston, 6431 Fannin Street, MSB 4.331, Houston, TX 77030. Tel.: +1 702 321 6559; fax: +1 713 566 4242.
E-mail address: Julie.L.Holihan@uth.tmc.edu (J.L. Holihan).

Introduction

Radiologic imaging is deeply integrated in medical care, and last year, 68 million Americans underwent a computed tomography (CT) scan.[1,2] Increasingly, CT scanning of the abdomen and pelvis is being used to diagnose and assess patients with a ventral hernia (VH).[3] However, a CT scan can both over- and under-diagnose hernias. Previous studies have shown poor agreement and poor interobserver reliability (κ) among radiologists reading CT scans for diagnosis of VHs (32% agreement, κ = 0.21, P < 0.001).[4]

Common reasons for disagreement of the presence or absence of a VH on CT scan have been identified, and recommendations to help improve agreement have been put forth. Recommendations include developing a standardized definition for VH versus mesh or tissue eventration, developing a systematic method for reviewing the abdominal wall for a hernia, and standardizing the communication between surgeons and radiologists.[4] Scans were reread immediately after the development of these recommendations, and agreement marginally improved (40% agreement, κ = 0.34, P < 0.001). As a quality improvement (QI) project, these recommendations have been incorporated into practice by the radiology department over the past year. However, the long-term effect of the implementation of these recommendations has not been assessed.

The aim of this QI project was to determine whether the incorporation of consensus recommendations developed by the surgeons and radiologists improves agreement among reviewers of CT scans in diagnosing a VH.

Methods

After the Institutional Review Board's approval at the University of Texas Health Science Center at Houston, all patients with surgery clinic appointments at a single institution from October to December 2016 were reviewed. A QI project was initiated in 2015 following the Standards for QUality Improvement reporting Excellence, version 2.0 guidelines.[5] A group of surgeons and radiologists developed and published consensus recommendations in May 2015.[4] These recommendations were introduced and disseminated throughout both departments and were discussed regularly at departmental meetings in an attempt to incorporate them into daily practice. This study was conducted 1 y after the beginning of this QI initiative.

Patients with a CT scan in the last 1 y and no intervening abdominal surgery were eligible to be included in the study. After obtaining informed consent, patients underwent a standardized examination in clinic for the presence or absence of a VH by a trained hernia surgeon. The clinician rated each patient as having no clinical hernia, indeterminate hernia, or clinical hernia. In general, examinations were classified as indeterminate likelihood if obesity precluded what the clinician perceived to be an adequate physical examination or if no hernia was palpated, but patients complained of localized pain or discomfort in the region in question. Baseline patient and hernia characteristics were recorded including gender, age, body mass index, American Society of Anesthesiologists score, smoking status, diabetes mellitus, hernia type (primary versus incisional hernia), and hernia width. P-values were calculated comparing groups based on clinical likelihood of hernia using the appropriate statistical test.

In addition, patient-centered outcomes were determined using patient surveys given at each clinic visit. Patients were asked to rate satisfaction and cosmetic satisfaction with their abdomen (1-10 Likert scale: 1 = dissatisfied and 10 = satisfied), abdominal pain (visual analog scale: 1 = no pain and 10 = severe pain), and patient function (using a modified activities assessment scale: 1 = poor function and 10 = good function). An overall function score was calculated as the sum of all of the function questions normalized to 100 (1 = poor function and 100 = good function).[6-8] Patient-centered outcomes were correlated to the number of radiologists who identified a hernia in each scan.

For each enrolled patient, seven independent radiologists reviewed the CT scans and were instructed to assess the scans for the presence or absence of a VH. They were given information on the clinical likelihood of a hernia. Subsequently, all seven radiologists and a hernia surgeon met to discuss the CT scans that had less than 100% agreement. Scans were reviewed and discussed as a group, and the radiologists revoted on the presence or absence of a VH. The consensus opinion of this group was considered the "gold standard" for the presence or absence of a VH; scans with no consensus remained documented as "no consensus".

Statistics

Percent agreement was calculated among all seven radiologists. In addition, the kappa statistic, κ, was used to determine interobserver reliability in identifying a VH on CT scan. K determines the magnitude of agreement between observers beyond that expected by chance alone.[9] A κ = 0 indicates
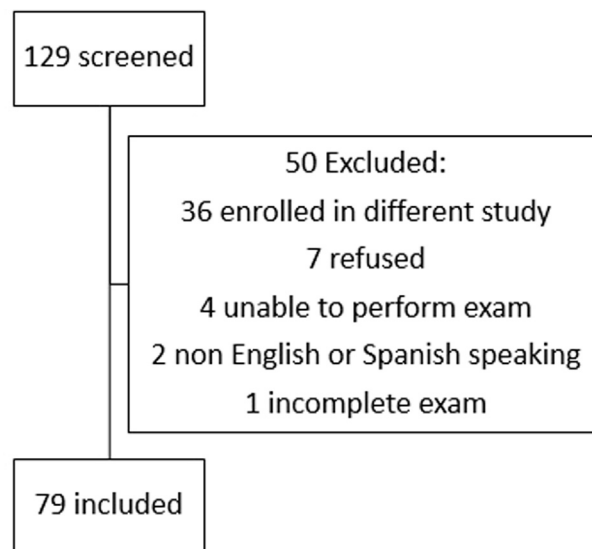

Fig. 1 — Flow sheet of included patients.