

Canadian Association of General Surgeons and American College of Surgeons Evidence-Based Reviews in Surgery. 26.

Watchful waiting versus repair of inguinal hernia in minimally symptomatic men

Jeffrey Barkun, MD; Amy Neville, MD; G. William N. Fitzgerald, MD; Demetrius Litwin, MD; for the Evidence-Based Reviews in Surgery Group*

CAGS Evidence Based Reviews in Surgery

The term "evidence-based medicine" was first coined by Sackett and colleagues as "the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients."¹ The key to practising evidence-based medicine is applying the best current knowledge to decisions in individual patients. Medical knowledge is continually and rapidly expanding, and it is impossible for an individual clinician to read all the medical literature. For clinicians to practise evidence-based medicine, they must have the skills to read and interpret the medical literature so that they can determine the validity, reliability, credibility and utility of individual articles. These skills are known as critical appraisal skills. Generally, critical appraisal requires that the clinician have some knowledge of biostatistics, clinical epidemiology, decision analysis and economics, as well as clinical knowledge.

The Canadian Association of General Surgeons (CAGS) and the American College of Surgeons (ACS) jointly sponsor a program entitled "Evidence-Based Reviews in Surgery (EBRS)," which is supported by an educational grant from ETHICON and ETHICON ENDO SURGERY, both units of Johnson & Johnson Medical Products, a division of Johnson & Johnson, and ETHICON INC. and ETHICON ENDO-SURGERY, INC. divisions of Johnson & Johnson Inc. The primary objective of this initiative is to help practising surgeons improve their critical appraisal skills. During the academic year, 8 clinical articles are chosen for review and discussion. They are selected not only for their clinical relevance to general surgeons, but also because they cover a spectrum of issues important to surgeons; for example, causation or risk factors for disease, natural history or prognosis of disease, how to quantify disease (measurement issues), diagnostic tests and the early diagnosis of disease, and the effectiveness of treatment. A methodologic article is supplied that guides the reader in critical appraisal of the clinical article. Both methodologic and clinical reviews of the article are performed by experts in the relevant areas and posted on the EBRS website. As well, a listserv discussion is held where participants can discuss the monthly article. Members of CAGS and ACS can access Evidence-Based Reviews in Surgery through the CAGS website (www.cags-accg.ca) or the ACS website (www.facs.org). All journal articles and reviews are available electronically through the EBRS website. We also have a library of past articles and reviews that can be accessed at any time. Surgeons who participate in the monthly packages can obtain Royal College of Physicians and Surgeons of Canada Maintenance of Certification credits and/or continuing medical education credits for the current article only by reading the monthly articles, participating in the listserv discussion, completing the monthly online evaluation and answering the online multiple choice questionnaire. For further information about EBRS, the reader is directed to the CAGS or ACS website or should email the administrator, Marg McKenzie, at mmckenzie@mtsinai.on.ca.

In addition to making the reviews available through the CAGS and ACS websites, 4 of the reviews are published in condensed versions in the *Canadian Journal of Surgery* and 4 in the *Journal of the American College of Surgeons* each year. We hope readers will find EBRS useful in improving their critical appraisal skills and also in keeping abreast of new developments in general surgery. Comments regarding EBRS may also be directed to mmckenzie@mtsinai.on.ca.

Reference

1. Evidence-Based Medicine Working Group. Evidence-based medicine. *JAMA* 1992;268:2420-5.

*The CAGS Evidence-Based Reviews in Surgery Group comprises Drs. J.S.T. Barkun, K. Brasel, T.H. Cogbill, C. Suzanne Cutter, G.W.N. Fitzgerald, H.J.A. Hentleff, A.W. Kirkpatrick, S. Latosinsky, A. Maclean, T.M. Mastracci, R.S. McLeod, L.A. Neumayer, S. Smith and L.R. Temple and Ms. M.E. McKenzie.

Correspondence to: Ms. Marg McKenzie, RN, Administrative Coordinator, EBRS, Mount Sinai Hospital, L3-010, 60 Murray St., PO Box 23, Toronto ON M5T 3L9; fax 416 586-5932; mmckenzie@mtsinai.on.ca

Selected Article

Fitzgibbons RJ Jr, Hurder-Giobbie A, Gibbs JO, et al. Watchful waiting vs. repair of inguinal hernia in minimally symptomatic men. A randomized clinical trial. *JAMA* 2006;295:285–92.

Abstract

Objective: To compare pain and physical function in men with minimally symptomatic inguinal hernia with watchful waiting or surgical repair. **Design:** Randomized controlled trial. **Setting:** Five community and academic centres in Canada and the United States. **Patients:** A total of 724 men were randomly assigned to a watchful waiting group ($n = 366$) versus a standard Lichtenstein open tension-free repair group ($n = 358$). Random assignment of patients was stratified by the presence of primary or recurrent hernia, unilateral or bilateral hernia and study site. **Intervention:** Patients in the watchful waiting group received instructions to watch for hernia symptoms and to contact their physicians if problems developed. Follow-up for these patients occurred at 6 months and annually. Patients who received the standard open tension-free repair were followed up at 3 months, 6 months and annually. **Main outcome measure:** Primary outcomes were pain and discomfort that interfered with usual activities 2 years after enrollment and changes from baseline to 2 years in the physical component score (PCS) of the SF-36 Health Survey, version 2. Secondary outcomes included complications, patient-reported pain, functional status, activity levels and satisfaction with care. **Results:** Primary intention-to-treat outcomes were similar at 2 years for patients in the watchful waiting and surgical repair groups (pain-limiting activities 5.1% in the watchful waiting group v. 2.2% in the surgical repair group, $p = 0.52$; PCS improvement over baseline 0.29 points in the watchful waiting group v. 0.13 points in the surgical re-

pair group, $p = 0.79$). Twenty-three percent of patients assigned to the watchful waiting group crossed over to receive surgical repair (increased pain was the most common reason); 17% of patients assigned to the surgical repair group crossed over to watchful waiting. The occurrence of postoperative hernia-related complications was similar among patients who received surgical repair as assigned and among patients in the watchful waiting group who crossed over to the surgical repair group. One watchful waiting patient (0.3%) experienced acute hernia incarceration without strangulation within 2 years. A second had acute incarceration with bowel obstruction at 4 years. The authors observed a frequency of 1.8 events per 1000 patient-years, including patients followed for up to 4.5 years. **Conclusion:** Watchful waiting is a safe and acceptable option for men with asymptomatic or minimally symptomatic inguinal hernias. Delaying surgery until symptoms increase is safe because acute hernia incarcerations occur rarely.

Commentary

Groin hernias are a common problem with an estimated prevalence of 3%–4% in the population¹ and a lifetime incidence of 27% in men.² Hernia repair is the third most commonly performed surgical procedure in Canada, with about 50 000 hernia repairs being performed yearly.³ However, surprisingly little is known about the natural history of untreated inguinal hernias, and questions remain regarding the optimal management of patients with asymptomatic or minimally symptomatic hernias.

Traditionally, it has been taught that all inguinal hernias should be repaired. One of the primary indications for hernia repair is relief of hernia-related symptoms such as pain. Although unproven, another common indication for repair is to prevent progression of the hernia to a size that could make repair more difficult. This is less relevant in the era of mesh-

based repairs, where the integrity of hernia repair depends more on the prosthetic mesh than on the quality of the patient's tissues.⁴ Finally, it is taught that hernia repair should be performed to prevent the occurrence of complications such as acute incarceration or intestinal strangulation.⁵

The benefits of hernia repair must be balanced with the risk of complications, both immediate and long-term, as well as socioeconomic considerations.⁶ Although generally minor, postoperative complications occur in about 20% of open mesh repairs.⁷ Long-term complications such as chronic pain severe enough to interfere with everyday activities occur in about 10% of patients after repair.⁸ Finally, socioeconomic considerations such as the cost of operation and caregiver burden cannot be ignored. Time lost from work is also an important consideration, because convalescence time following hernia repair is variable and is affected by a number of factors, including occupation, the possibility of workman's compensation, etc.⁹

With this in mind, Fitzgibbons and colleagues¹⁰ compared men with asymptomatic or minimally symptomatic inguinal hernias treated with watchful waiting to standard open tension-free repair with mesh. The main outcomes were pain and discomfort interfering with usual activities; secondary outcomes included operative complications, change in activity levels and patient satisfaction. The study also sought to determine the natural history of minimally symptomatic untreated hernias and the associated risk of hernia accidents.

The authors conducted a multicentre, randomized controlled trial involving men aged 18 years and older who had inguinal hernias that were either asymptomatic or minimally symptomatic. A minimally symptomatic hernia was defined as "the absence of hernia-related pain or discomfort limiting usual activities." Of the 3074 men screened for the trial, about half were considered to be ineligible, and 50% of the eligible pa-

tients declined to participate. Thus 724 patients, representing about 25% of patients initially screened, were randomly assigned to the watchful waiting group or to the surgical repair group. This participation rate is consistent with expected standards for clinical trials; however, it is of concern whether the population included in the study is representative of the intended target population, questioning external validity.

As part of the randomization process, study participants were stratified by the presence of primary or recurrent hernia, unilateral or bilateral hernia and study site. In smaller randomized clinical trials, stratification of variables that are known to affect outcome is an important way to ensure a similar distribution of these variables across groups. Randomization alone may not guarantee equal representation in trials with small numbers of participants. In larger trials (those with more than 1000 patients), the larger sample allows for the balancing of effects attributable to these variables, and stratification becomes less important.¹¹ Assignment to the treatment group was determined by central computerized randomization in permuted blocks of 2, 4 and 6. Randomization in blocks ensures that patients are assigned in equal numbers to each treatment arm. Varying the block size prevents site investigators from being able to predict to which treatment arm successive patients will be assigned. Randomization in this trial was effective in creating groups with similar baseline characteristics. However, there were statistically significant differences between groups with respect to body mass index (BMI) (a difference of 1.2 m/kg²), 3 of the Activity Assessment Scale (AAS) scores, and in the proportion of patients with enlarging hernia in the 6 weeks preceding trial enrolment (greater proportion in the watchful waiting group). Although statistically significant, it is unlikely that the BMI and AAS differences were clinically meaningful. If these variables had

been considered to be clinically important, they would have been included in the stratification scheme at the outset of the trial.

Given the nature of the trial (watchful waiting v. surgical repair), blinding of investigators or participants was not feasible. The median follow-up time was 3.2 years. Long-term follow-up of patients in clinical trials is of key importance but often challenging. In this trial, follow-up was complete in 90% of patients, and the completeness of follow-up is a strength of the study.

The main outcomes of the trial were pain and physical function, and these were measured in 2 ways. First, pain and discomfort limiting usual activity was measured as a level-3 or 4 response on a 4-point Likert scale. Second, the SF-36 Health Survey, version 2, health-related quality of life form was used. The SF-36 is a questionnaire intended to measure some components of patient-perceived quality of life, and analysis of the SF-36 involves the tabulation of 2 standardized summary scores: the mental and physical component scores. A change in PCS was used as the other primary outcome measure for the study. Secondary outcomes in the trial included postoperative complications, hernia recurrence, incidence of hernia accidents, patient satisfaction (measured on a 5-point Likert scale) and crossover rates between groups.

The study was powered to detect a 10% difference in the proportion of patients who experienced pain that interfered with daily activities and an 8-point difference in the PCS with 90% power and a 2-sided type I error rate of 5%. The intention-to-treat analysis revealed no significant difference between groups with regards to pain limiting normal activities or change in PCS at 2 years. The intended power of 90% in this trial was higher than in most trials (80% power being more common), and the minimum detectable difference was set at 10%, which is quite low. Both of these are charac-

teristic of a noninferiority trial (a trial whose intention is to show that a new practice is no worse than an established practice). In such a context, unlike conventional superiority trials, the absence of a difference between groups suggests clinical end-point equivalence.

The operative complication rate was 21.7% in the surgical repair group and was not significantly different from that of patients who crossed over during the course of the trial and had surgery after a period of watchful waiting (27.9% complication rate). Most complications were minor, with the most common being wound hematoma, scrotal hematoma, urinary tract infection and wound infection. Life-threatening complications occurred in only 3 patients who underwent surgery. The recurrence rate at 2 years was 1% among patients in the surgical repair group and 2.3% among patients who were assigned to the watchful waiting group but crossed over to surgical repair. The difference in recurrence rates between these 2 groups was not statistically significant (although the trial design was not powered to detect this difference). The incidence of hernia accidents in the watchful waiting group was exceedingly rare with a rate of 0.6% overall or 1.8 events per 1000 patient-years.

Based on their findings, the authors concluded that watchful waiting is an acceptable option for men with minimally symptomatic inguinal hernias and that delaying surgical repair until symptoms increase is safe because acute hernia incarcerations occur rarely.

One of the challenges in adopting new information acquired from any clinical trial is ensuring that the findings apply to the patient population to which the clinical practice will be applied. The conclusion of this trial by Fitzgibbons and colleagues is heavily dependent on the definition of a "minimally symptomatic" hernia. However, no explicit definition of "minimally symptomatic" hernia was provided. At baseline, about 8% of the

study participants stated they had pain at rest and more than 15% had pain with exercise. It would be important to quantify how much pain these patients had to ensure a consistent definition of "minimally symptomatic" patients. To do this, the authors could have evaluated symptoms on a screening questionnaire or a baseline visual analogue scale and subsequently provided those tools to clinicians at large.

The study population also included patients recruited by radio advertising. It is important to note that these self-referred participants may be considerably different (with respect to symptomatology, symptom severity and health-related behaviour) than patients who seek medical attention and are referred for surgical care. A recent European review also cited the fact that a large proportion of patients (40%) had hernias that were not visible and in fact only palpable on impulse; this was considered to be a limitation to the external validity of this study.¹² However, the typical size of hernias upon presentation tends to be practice-specific and one could postulate that the previous comments may reflect a later pattern of referral seen within the National Health Service in the United Kingdom.

A considerably high crossover rate was observed in this trial and, as such, one must evaluate whether the crossover rate represents a threat to the internal validity of the study. The mean time to crossover was 27.3 months. Characteristics of patients who crossed over appeared to be relevant to clinical decision-making. Seventeen percent of patients assigned to the surgical repair group did not have surgery. These patients had a higher American Society of Anesthesiologists (ASA) classification and a higher prevalence of diabetes and hypertension, indicating that they were likely less healthy than those who had surgical repair as per randomization. After 4 years, nearly one-third (31%) of the watchful waiting group had crossed over to surgical repair. These crossover patients had higher levels of hernia-related pain and impaired physical

function at baseline. In the context of this trial, the high crossover rate observed from watchful waiting to surgical repair should be considered to be more characteristic of an outcome measure than a threat to internal validity. The patient characteristics identified may prove helpful in predicting patients for whom watchful waiting will eventually lead to surgery. Furthermore, one could argue that 69% avoided surgery over 4 years with the strategy of watchful waiting.

Several of the secondary outcomes evaluated in the trial were found to be significantly different between groups and actually favoured the surgery group. These outcomes included an improvement in the perception of pain unpleasantness and greater improvement in the ability to perform everyday activities in the surgical repair group. Although these findings are clinically interesting, they were designed to be secondary outcomes and should be seen as hypothesis-generating rather than confirmatory of a pre-hoc hypothesis. It is for this reason that the authors did not highlight these statistically significant findings in their conclusions.

An important contribution of this trial is our enhanced understanding of the natural history of minimally symptomatic hernias and hernia accident rates in contemporary times. This aspect of the trial is key in answering the most important question that arises when considering a change in clinical practice: Do the benefits outweigh the potential harms and costs? The aim of this trial was to determine whether immediate surgery can be delayed or omitted altogether in patients with asymptomatic or minimally symptomatic groin hernias. The relatively short period of observation of the trial allowed for only a partial answer. This trial showed that there is no significant difference in pain or patient satisfaction between immediate surgical repair and watchful waiting and that the watchful waiting group was not exposed to an inordinately high risk of hernia accident. More-

over, the ultimate development of surgical complications was not increased by delaying surgery.

Therefore, it appears that given patient preference, observation is indeed a feasible and valid alternative to mandatory surgery in the short-term.

Competing interests: None declared.

References

1. Hay JM, Boudet MJ, Fingerhut A, et al. Shouldice inguinal hernia repair in the male adult: the gold standard? A multicenter controlled trial in 1578 patients. *Ann Surg* 1995;222:719-27.
2. Kingsnorth A, LeBlanc K. Hernias: inguinal and incisional. *Lancet* 2003;362:1561-71.
3. *Surgical procedures and treatments 1992-93*. Ottawa: Statistics Canada; 1995. Cat no 82-217.
4. Turaga K, Fitzgibbons RJ, Puri V. Inguinal hernias: Should we repair? *Surg Clin North Am* 2008;88:127-38.
5. Eubanks WS. Hernias. In: Townsend CM, editor. *Sabiston textbook of surgery. The biological basis of modern surgical practice*. 16th ed. Philadelphia: WB Saunders; 2001. p. 783-891.
6. Fitzgibbons RJ Jr, Jonasson O, Gibbs JO, et al. The development of a clinical trial to determine if watchful waiting is an acceptable alternative to routine herniorrhaphy for patients with minimal or no hernia symptoms. *J Am Coll Surg* 2003;196:737-42.
7. Payne JH, Grininger LM, Izawa MT, et al. Laparoscopic or open inguinal herniorrhaphy? A randomized prospective trial. *Arch Surg* 1994;129:973-81.
8. Cunningham J, Temple WJ, Mitchell P, et al. Cooperative hernia study. Pain in the post-repair patient. *Ann Surg* 1996;224:598-602.
9. Barkun JS, Keyser EJ, Wexler MJ, et al. Short-term outcomes in open vs. laparoscopic herniorrhaphy: confounding impact of worker's compensation on convalescence. *J Gastrointest Surg* 1999;3:575-82.
10. Fitzgibbons RJ Jr, Hurder-Giobbie A, Gibbs JO, et al. Watchful waiting vs. repair of inguinal hernia in minimally symptomatic men. A randomized clinical trial. *JAMA* 2006;295:285-92.
11. Kjaergard LL, Villumsen J, Glud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. *Ann Intern Med* 2001;135:982-9.
12. O'Dwyer PJ, Chung L. Watchful waiting was as safe as surgical repair for minimally symptomatic inguinal hernias. *Evid Based Med* 2006;11:73-4.