

Editorial

The compelling rationale for less preoperative testing

Michael F. Roizen MD

The study by Charpak *et al.*¹ in this issue is one of more than twenty studies (summarized in reference 2) demonstrating that the number of preoperative tests can be reduced, without adversely affecting patient care. While Charpak *et al.* focused on one specific test, the chest x-ray, the overall conclusion of studies either limited to one test, or including a battery of tests is the same – approximately 60 per cent of preoperative testing could be eliminated without adversely affecting patient care. One therefore has to ask two questions: Why isn't this being done? Why are we still seeing studies like that of Charpak *et al.* which appear to reconfirm what is already known?

Because of familiarity, I will address this topic in terms of health care and medical practice patterns in the United States which, although on a different financial basis than in Canada, nevertheless illustrate the general problems of implementing a strategy of lesser preoperative testing.

Let us first examine the subject in two ways. (1) Are there advantages to ordering fewer tests – advantages to society, to patients, and/or to providers? (2) If there are advantages, why is it taking us so long to implement these strategies? Is it the milieu in which testing is ordered or is it because the advantages to providers aren't as great as the advantages to society?

First, are there advantages to ordering fewer tests? The data from many published studies are fairly consistent and show that there are advantages to society in providing less testing; that is, cost savings and perhaps improvement of patient care by reducing risk to patients. This latter statement may seem surprising. In medicine, one almost always assumes a priori that doing more testing would be better, and produce more of a benefit. But this is not necessarily true, especially if the extra testing yields false-positive results or clinically unimportant borderline positive results which lead to both delays in the surgical

schedule and, more importantly, increased risk to the patient as more testing is done and more judgements are based on the false-positive or borderline tests. Increased risk created by extra testing is actually one of the new risks which have emerged during the last two decades.

The Kaiser organization, for example, instituted multiphasic screening with 16 tests in the mid 1960's. While previous to that tests were ordered singly or in small groups to confirm clinical impressions, the advent of automated screening batteries and of the "executive physical" led to increased numbers of tests. The health maintenance organizations in the United States were one of the main proponents of screening tests. The rationale appeared good and very logical. You could screen cheaply and spot abnormalities before they led to symptoms and thus prevent illness from occurring. It has not worked out this way and, in fact, extra tests have simply posed extra risk to the patient, because of the iatrogenic disease caused by the pursuit of and treatment of borderline abnormal and false-positive results. In fact, the Kaiser organization has dropped the majority of the batteries of screening tests, and certainly for preoperative care. This is not to say that some screening is not indicated – some obviously is, such as the mammogram for females over the age of 50, the stool for occult blood test for patients over the age of 40, the "pap smear," etc. But random screening, without being selective, may not be more of a benefit.

Few studies have examined the adverse results to patients of increased tests and the follow-up of false-positive tests. We examined this question retrospectively for chest x-rays.³ In this study, 386 extra chest x-rays were done that were not indicated, in a population of 606 patients. One problem, an elevated hemidiaphragm and probable phrenic nerve palsy was found that may have improved care for that patient (1/386). In addition, three lung shadows were found, which resulted in three sets of invasive tests and one thoracotomy, without any disease being discovered and with considerable morbidity for the three patients, including one pneumothorax and four months of disability.

Professor and Chairperson, Department of Anesthesia and Critical Care, The University of Chicago, 5841 South Maryland Avenue, Chicago, Illinois, 60637.

In another study,⁴ a series of eight years of routine testing of 121 nursing home patients, who were on the average 89 years of age, resulted in no benefit being derived from 74 per cent of the tests. This result could have been predicted before the testing began. In fact, one patient was seriously and adversely hurt by the extra testing and treatments it engendered. Thus, I perceive that selective testing of the type done before automated screening batteries is better for patient care, as well as saving society approximately 50 to 60 per cent of the expense of this testing.

Blue Cross/Blue Shield has estimated that more than 30 billion dollars is spent in North America on preoperative testing and the subsequent follow-up evaluations, because of those tests. They have further estimated that 18 billion dollars, or 60 per cent of this, could be saved by appropriate ordering of tests, based on the history or on findings of the physical examination. Since selectively ordering tests is better for patients and saves money, we conclude that providers must perceive a benefit from the current extra testing for them to continue this adverse and more costly pattern of care. Perhaps these perceived benefits are the medico/legal protection or an efficiency of scheduling of their lives. I believe both are perceived, but neither are true. Let me explain the basis for this belief.

Are extra tests medico/legal protection? The answer is a resounding "no." There are now a series of reports (reviewed in references 2 and 3) that show that between 30 and 60 per cent of all unexpected abnormalities found on preoperative laboratory tests are not noted preoperatively. Clearly this is the case with many preoperative x-rays, where the report is not on the chart before the start of the anaesthetic. An abnormality not appropriately pursued is a greater liability risk than is an abnormality that is not detected preoperatively.⁵ It appears that extra testing, testing that is not indicated by history, is an extra risk to patients and is an extra medico/legal risk to physicians.

Thus it seems that the only advantage to continuing the present system of the routine ordering of batteries of screening tests is the perceived efficiency for operating room schedules by surgeons attempting to appease anaesthesiologists. What could solve the problem is a system (but still with surgeon and anaesthesiologist intervention based on clinical judgement) where the anaesthesiologist agrees that certain conditions indicate specific tests and certain other conditions do not require tests. If one could automate the system the surgeons could institute it in the same time that they would normally take to order the current battery of tests.

The study by Charpak *et al.*,¹ where the surgeons and anaesthesiologists agreed for research purposes on set conditions for ordering chest x-rays is an example of such a system. The set conditions were: any lung disease, any

cardiovascular disease, actual malignant disease, major surgical emergencies, current smoking history in patients ≥ 50 years old, immunodepression, and lack of a prior health examination in immigrants. Even so, in the study by Charpak *et al.*, 271 chest x-rays were ordered which were not recommended and 596 were not ordered although they were recommended, out of a total of 1426 chest x-rays that should have been ordered in this group of 3,849 patients. If there were this many errors in the ordering of one laboratory test, inclusion of more tests would certainly generate more errors.

What appears needed to reduce the error rate for ordering is a system which generates the medical history and from the history suggests tests for the physician to order. We tested a protocol in which the results of a history filled out by the patient on six pages of questions suggested which laboratory tests should be ordered.^{3,6} This system only worked when a health research nurse was present to integrate the results of the history into agreed-upon indications for each test. It appears, then, that an automated system is necessary, since hiring a health nurse in every location where physicians order preoperative tests is likely to be more expensive than the testing itself. We tried a computer-based system based on a MacIntosh computer, but this did not work, because many patients are not "computer literate." We are now trying another system which will give a "health quiz" to patients on a simple four-button computer machine like the child's game "Donkey Kong, Jr." The surgeon or anaesthesiologist will have this device in his/her office, where it can be used by the patient. The device asks patients yes/no questions and then prints their answers, as well as the suggested laboratory tests, based on the agreed upon indications and the patients' answers. The physician, surgeon, or anaesthesiologist can over-ride or add to the suggested tests, before the preoperative testing is actually ordered.

Thus, whether or not an easy implementation mechanism is available, it appears that less testing is in order. Less testing will reduce the risk to the patient, reduce medico/legal risks for the physician, lessen costs to society, and probably make operating room schedules more efficient, by not creating the need to pursue falsely positive or slightly abnormal results with which the patient normally lives. I believe that the time has come when the studies of Charpak *et al.*, and others, will beget benefit to society and that less but more appropriate testing will be a reality for the future.

La rationnelle pour la diminution du nombre de tests préopératoires

L'étude de Charpak *et al.*¹ publiée dans ce numéro est une parmi plus d'une vingtaine d'autres études (résumé dans la référence 2) démontrant que le nombre de tests préopératoires peut être réduit sans nuire au soin des patients. Alors que Charpak *et al.* se sont attardés sur un test spécifique, le rayon-x pulmonaire, la conclusion générale d'autres études, qu'elles soient limitées à un test ou qu'elles incluent une batterie de tests reste la même: approximativement 60 pour cent des tests préopératoires peuvent être éliminés sans préjudice aux patients. On doit donc se poser deux questions: d'abord pourquoi ceci n'est pas encore fait? Pourquoi voit-on encore des études comme de Charpak *et al.* qui ne font que reconfirmer ce qu'on sait déjà?

Parce qu'il m'est plus familier, je vais adresser ce sujet selon les normes de soin de santé et le profil de pratique médicale connus aux Etats-Unis. Malgré qu'ils soient différents financièrement du système Canadien ces normes de pratique illustrent néanmoins les problèmes généraux qui empêchent l'implantation d'une stratégie pour diminuer les tests préopératoires.

Exammons d'abord le sujet de deux façons. 1) Y-a-t-il des avantages à la société, aux patients, et/ou à aux pourvoyeurs à diminuer le nombre de ces tests? 2) En l'absence d'avantages, pourquoi l'application d'une telle stratégie prend-t-elle du temps? Cela dépendrait-il du milieu ou serait-ce parce que les avantages que cette stratégie amène aux pourvoyeurs ne seraient pas aussi grands que les avantages à la société?

Premièrement y-a-t-il des avantages à diminuer le nombre de tests? Les données publiées dans plusieurs études sont raisonnablement consistantes et démontrent des avantages à la société quand on adopte une pareille stratégie que ce soit par la diminution des coûts ou peut être par la diminution des risques que peuvent amener ces tests. Ce dernier point peut sembler surprenant. En médecine on peut toujours assumer à priori qu'en faisant plus de tests on aboutira à un plus grand bénéfice au patient. Ceci n'est cependant pas toujours vrai. Spécialement si le surplus de tests aboutit à des résultats faussement positifs ou des résultats positifs mais sans importance clinique. Le retard dans la cédule chirurgicale et le risque accru que cela amène sera basé sur des tests faussement positifs. Un risque augmenté créé par des tests

supplémentaires est en fait un des nouveaux risques qui a vu le jour dans les deux dernières décennies.

L'organisation Kaiser, par exemple, a institué un système de dépistage multiphasique comprenant 16 tests vers le milieu des années 1960. Alors qu'auparavant les tests étaient demandés à l'unité ou par petits groupes afin de confirmer les impressions cliniques, l'établissement de ce système de dépistage automatisé a amené une surmultiplication des tests accomplis. Les organisations de santé des Etats-Unis ont favorisé ces tests de dépistage. La rationnelle apparaissait bonne et logique. On pouvait dépister à bas prix, détecter des anomalies avant le déclenchement des symptômes et prévenir ainsi des maladies. Cependant le résultat était différent car en fait les tests supplémentaires ont créé un risque supplémentaire iatrogène au patient par le traitement des patients qui présentaient des tests à la limite de la normale ou des résultats faussement positifs. En fait l'organisation Kaiser a abandonné la majorité des batteries de tests de dépistage et plusieurs des tests préopératoires. Ceci ne veut pas dire que les tests de dépistage ne sont pas indiqués: quelques-uns de toute évidence le sont comme c'est le cas de la mammographie chez les femmes dont l'âge est supérieur à 50 ans, le test de dépistage du sang dans les selles chez les patients âgés de plus de 40 ans, le "pap smear", etc. Un dépistage randomisé sans être sélectif peut ne pas être bénéfique.

Peu d'études se sont attardées sur l'effet néfaste d'une augmentation des tests et la poursuite des tests faux-positifs. On a examiné cette question rétrospectivement pour les rayons-x pulmonaires.³ Dans cette étude, 386 rayons-x pulmonaires additionnels ont été demandés sans indication spécifique pour une population de 606 patients. Un seul problème pouvant améliorer les soins du patient fut détecté soit une surélévation de diaphragme suggérant la possibilité d'une paralysie du nerf phrénique (1/386). En plus, trois opacités pulmonaires furent détectées provoquant ainsi d'autres tests invasifs et créant une morbidité accrue dont une thoracotomie qui fut négative, un pneumothorax et une incapacité de quatre mois.

Dans une autre étude,⁴ où des tests de routine étaient faits durant huit ans dans 121 centres de soins de patients âgés en moyenne de 89 ans, on a démontré qu'aucun bénéfice n'a été tiré à partir de 74 pour cent des tests. Ce résultat pouvait être prédit avant le début de l'étude. En fait, les tests supplémentaires et les traitements qu'ils ont engendrés ont nui sérieusement à un patient du groupe. Ainsi, je crois que demander des tests sélectifs du type qu'on faisant avant l'instauration des batteries de dépistage automatisé serait meilleur pour les soins des patients et économiserait à la société approximativement 50 à 60 pour cent des dépenses.

Les compagnies d'assurance Blue Cross/Blue Shield

ont estimé que plus de 30 billions de dollars sont dépensés dans le continent Nord Américain sur des tests préopératoires et la poursuite des résultats de ces tests. Ils ont de plus estimé que 18 billions de dollars, ou 60 pour cent de ceci pourraient être sauvés par une demande appropriée des tests basés sur l'histoire ou sur les résultats de l'examen physique. Etant donné qu'une demande sélective des tests est meilleure pour les patients et peut sauver de l'argent on conclut que les pourvoeureurs de ces tests doivent y percevoir des bénéfices personnels afin de continuer ce programme dispendieux. Il est possible que les bénéfices perçus soient la protection médico-légale ou une meilleure efficacité dans la cédule de leur vie. Je crois que ces deux points sont perçus mais qu'aucun d'eux n'est vrai. Laissez-moi expliquer mon point de vue.

Est-ce que des tests supplémentaires amènent une plus grande protection médico-légale? La réponse est un "non" éclatant. Il y a actuellement une série d'études (révisées dans les références 2 et 3) qui démontrent qu'entre 30 et 60 pour cent des anomalies découvertes fortuitement par les tests préopératoires ne sont pas notées dans cette période. Ceci est nettement le cas des rayons-x pulmonaires préopératoires ou les résultats n'étaient pas dans le dossier lors du début de l'anesthésie. Une anomalie non explorée adéquatement représente une plus grande responsabilité légale qu'une anomalie non détectée en période préopératoire.⁵ Il apparaît ainsi qu'un surplus de test, qu'un test non indiqué par l'histoire, représente un risque supplémentaire aux patients ainsi qu'un plus grand risque médico-légal pour le médecin.

Ainsi le seul avantage à continuer le présent système de dépistage de routine serait de bâtrir une cédule opératoire efficace par des chirurgiens tentant d'appaiser les anesthésiologistes. Ce qui pourra résoudre le problème est un système (avec l'accord du chirurgien et de l'anesthésiologue et basé sur le jugement clinique) où l'anesthésiologue admet que certaines conditions demandent des tests spécifiques et d'autres conditions n'en requièrent point. Si quelqu'un pouvait automatiser ce système les chirurgiens pourraient l'instaurer de la même façon qu'ils demandent actuellement les batteries de tests.

L'étude de Charpak *et al.*¹ où les chirurgiens et les anesthésiologistes se sont mis d'accord pour fin d'étude sur des conditions précises où le rayon-x pulmonaire est requis est un exemple d'un tel système. Les conditions pré-établies étaient: toute lésion pulmonaire, toute atteinte cardio-vasculaire, toute maladie maligne, les urgences chirurgicales majeures, une histoire de tabagisme chez des patients âgés de 50 ans et plus, un état immunodépressif, et l'absence d'examen médical préalable chez les immigrants. Dans l'étude de Charpak *et al.*, 271 rayons-x pulmonaires étaient demandés sans indication et 596 n'étaient pas demandés même s'ils étaient

indiqués pour un total de 1,426 rayons-x pulmonaires qui auraient dûs être demandés dans ce groupe de 3,849 patients. S'il y avait autant d'erreurs dans la demande d'un seul test de laboratoire, l'inclusion d'autres tests aura certainement généré d'autres problème.

Ce qui apparaît nécessaire afin de diminuer le taux d'erreurs dans la demande des tests serait un système basé sur une histoire médicale à partir de laquelle on suggère des tests. On a essayé un protocole où le résultat de l'histoire rempli par le patient sur six pages de questionnaire suggère lequel des tests de laboratoire doivent être demandés.^{3,6} Ce système fonctionnait uniquement quand une infirmière de recherche était présente afin d'intégrer les résultats de l'histoire dans les indications préalablement suggérées pour chaque test. Il est évident qu'un système automatisé serait nécessaire car l'embauche d'une infirmière serait plus coûteux que le test lui-même. On a essayé un système programmé sur MacIntosh sans succès car plusieurs patients n'étaient pas familiers avec les ordinateurs. Actuellement on essaie un autre système qui fournit au patient un "quiz de santé" sur un simple ordinateur à quatre boutons identiques au jeu de "Donkey Kong, Jr". Le chirurgien ou l'anesthésiste aurait cet appareil dans leur bureau. L'appareil demande au patient des questions dont les réponses sont des simples "oui" ou "non" et imprime leurs réponses ainsi que les tests de laboratoire suggérés se basant sur les indications pré-établies. Le médecin, le chirurgien, ou anesthésiologue peut cependant passer outre ou ajouter aux tests suggérés par ce système avant que les tests préopératoires soient demandés.

Ainsi qu'un mécanisme d'application facile soit disponible ou pas il apparaît qu'on a besoin de diminuer le nombre de test. Moins de test réduit le risque au patient, réduit les risques médico-légaux pour le médecin, réduit les coûts à la société et rend probablement la cédule opératoire plus efficace. Je pense que le temps est arrivé pour que les travaux comme ceux de Charpak *et al.* amènent dans le futur un bénéfice à la société en diminuant le nombre des tests préopératoires et en les rendant plus appropriés.

References

- 1 Charpak Y, Blery C, Chastang C, Szatan M, Fourgeaux B. Prospective assessment of a protocol for selective ordering of preoperative chest x-rays. *Can J Anaesth* 1988; 35: 259-64.
- 2 Roizen MF. Routine preoperative evaluation. In: *Anesthesia*. Miller RD (Ed), 2nd edition, 1: 225-54. New York, Churchill Livingstone, 1986.
- 3 Roizen MF, Kaplan EB, Schreider BD, Lichter JL, Orkin FK. The relative roles of the history and physical examination and laboratory testing in preoperative evaluation for outpatient surgery: the "Starling Curve" of preoperative laboratory testing. *Anesthesiology Clinics of North America* 1987; 5(1): 15-34.
- 4 Levinstein MR, Ouslander JG, Rubenstein LZ, Forsythe SB. Yield of routine annual laboratory tests in a skilled nursing home population. *JAMA* 1987; 258: 1909.
- 5 Robertson WM. *Medical malpractice: a preventive approach*. Seattle, University of Washington Press, 1985.
- 6 Kaplan EB, Sheiner LB, Boeckmann MS et al. The usefulness of preoperative laboratory screening. *JAMA* 1985; 253: 3576-81.