Author’s response to reviews

Title: A survey of patient preference for colorectal cancer screening technique

Authors:

Prof Richard L Nelson MD (altohorn@uic.edu)
Prof Alan Schwartz Ph.D. (alansz@uic.edu)

Version: 2 Date: 7 Mar 2004

PDF covering letter
Thanks to Dr. Renehan for detailed and challenging comments. But, to the comments:

1 I would argue that we are comparing like with like:
   All 4 modalities are recommended by the ACS, NCI and USPSTF as valid screening procedures for colorectal cancer.
   I agree emphatically that screening for precancers has the greater potential for disease control and mortality reduction. However published guidelines still state that cancer detection is the goal of screening. Nevertheless, all 4 are capable of detecting precancers as well as established malignancies (see: Mandel JS, Church TR, Bond JH, Ederer MA, et al. The effect of fecal occult blood screening on the incidence of colorectal cancer. N Eng. J Med. 2000;343:1603-7), even if by indirect means. Each is a point on a path, with the end of the path being colonoscopy, but a single path nevertheless, and therefore like to like. And as stated, the primary challenge is to get more of the public on to the path, thus the need to find out why so many do not attend screening. A shorter form of this argument is already in the text.

2 Although we used non-parametric (rank) statistics, we chose to present mean ranks as an aid to the reader. Mean ranks are commonly presented when tests on distributions of ranks are performed, as they are more compact than presenting complete frequency tables and convey important information about the distributions, whether in injury severity scores such as the PATI, illness scores such as APACHE or in pain scores.

3 Small but adequate. No mention is made of our power calculation. We based our sample size on the size required for 80% power to detect a one-half standard deviation difference within subject using a paired t-test (34), and then conservatively assumed we would need 150% of that sample size (51). Would a bigger population have yielded more valid results? We have no evidence to suggest that. Highly selected? Yes and especially for the two specified characteristics in the text: non-participants in colorectal screening or diagnostics (this is the major flaw of all previous reports) or hospitalized patients (which creates its own set of biases). The age range was also intentionally chosen to get impressions before the screening decision had been presented to the participant by their physician, and also to get a broader age range than had been previously reported (in which most individuals were very old, in some cases well beyond ages where one might screen). There is an analysis that compares reported preferences in different age groups and genders in the text. Surveys of differing populations would be of interest.

To specific points:

1 OK

2 Reference 4 already given, and text and references are added to contrast with breast cancer screening.

3 True, but it may be extrapolated to sigmoidoscopy or barium enema. The problem is the practitioner’s expertise in interpreting the results of any of these tests.
4 True, but not yet the mortality data that exist for FOB. These are valuable contributions, but not directly relevant to this study, until their efforts might merge with this to develop more effective means of drawing non-participants into screening.

5 The letter was written by Dr. Nelson. The description of the study was by a clinic nurse mentioned in the acknowledgments: Ms. Campbell. Bias is a huge issue in any such survey and clearly present in all previous publication on this topic, in which participants were subjected to lengthy descriptions of the accuracy and “value” of each screening test. It was our impression that the choice might rest more on aesthetics than statistics, and so the actual preparation and conduct of the exam, I hope described objectively, formed the substance of our description. The exam instrument has been included with the report. It was our intent to distribute the survey in a location in which bias would also be less prevalent, compared for instance to a gastroenterologist’s office or radiology suite (where some previous surveys have been administered).

6 We are after the naïve impressions of a public that will choose whether or not to participate in screening, and, as above, like is being compared to like. Nor do I think are all the potential dangers of any screening test fully determined.

7 The age discussion is above. Why would we for instance want to survey a geriatric population, as others have?

8 OK

9 An Americanism: physicians being a generic term covering all specialties, and nurse endoscopists are still rare in this country.

10 I have never heard the term “buy into” used in a Wall Street context. It does connote in the U.S. voluntary participation in a program or endeavor in which the participant is expected to incur some cost, in pain, resources, inconvenience, etc. We could have said this differently but not as succinctly.

11 I suppose. The main reason to bring up gene based stool tests is that, like FOB, no bowel preparation is needed. Our anecdotal impression, and that of many others is that the primary objection the public has to other screening modalities is not the pain or embarrassment of the actual procedure, but the bowel prep the night before, which one would of course have to take for CT colography too.

Again, may thanks,

Rick Nelson

Although we used non-parametric (rank) statistics, we chose to present mean ranks as an aid to the reader. Mean ranks are commonly presented when tests on distributions of ranks are performed, as they are more compact than presenting complete frequency tables and convey important information about the distributions.

I wouldn't say your bit about averages and sd's in general, though.
Small but adequate. No mention is made of our power calculation.

Well, we didn't really give much detail on the power calculation, because there isn't one to give; power methods for our test statistics are not widely available, so we based our sample size on the size required for 80% power to detect a one-half standard deviation difference within subject using a paired t-test (34), and then conservatively assumed we would need 150% of that sample size (51). Maybe add that in?

I concur with your other responses.