Reviewer's report

Title: A review of randomized controlled trials evaluating the effectiveness of handheld computers for data collection.

Version: 1 Date: 10 July 2005

Reviewer: Jochen R Moehr

Reviewer's report:

This paper is of interest and should be published with extensive independent comments. It is of interest not for the reasons intended by the authors but because it is a very good illustration of the limits if not dangers of applying the RCT approach to the evaluation of information systems in general and particularly in health.

The reasons for this have been analyzed for over a decade, e.g., by:


Moehr, J., R., Anglin, C., Schaafsma, J., Pantazi, S., Grimm N.: Lest Formalisms Impede Insight and
Briefly put, information systems are not constant as e.g., drugs but are complex very invasive systems that typically change constantly over time and as they are adapted to various environments of use. Their effectiveness is highly dependent on many variables of the user population, circumstances and purposes of use, software and hardware technology employed, version used and state of development, etc. Therefore unexpected effects and adverse effects are common and have to be watched for. These can typically not be anticipated as required by the design of an RCT type approach. Therefore the information yielded by RCT type approaches to evaluation is usually lacking important and necessary detail and insights, no matter how well the anticipated effects are corroborated or refuted by the RCT design. The rest is usually not of much use for the advancement of knowledge or of health information systems, except perhaps in the rare cases of highly contained and highly specific devices, such as imaging devices or devices for biosignal interpretation. For a more comprehensive outline of the shortcomings of RCT in this context compare, e.g., Moehr 2002, and Ammenwerth 2004 above.

These shortcomings are well illustrated in the current paper by S. J. Lane et al:

The authors used an RCT approach to assess the usefulness of hand held computers over conventional paper-based approaches for the recording of scientific data in a haemophilia clinic. Probably prodded by this work of their own, they decided to do a systematic review of similar comparisons of hand held computers and conventional documentation practices used by patients and other non research personnel for collection of scientific data. They used the standard approach for systematic reviews and searched appropriate data bases with defined searches. They came up with 201 studies, which were assessed according to defined inclusion and exclusion criteria. Both sets of criteria are summarized in a Figure 1 (of which, however, only the inclusion criteria are visible in the copy available to this reviewer).

The first inclusion criterion is RCT Study design. If the authors were familiar with any of the literature cited above, it might have dawned on them that this is a perfect way to exclude the more relevant and insightful papers from consideration. No analysis of the 193 excluded articles beyond a very high level categorization, also given in Figure 1 is provided.

Another inclusion criterion is Subjects [are] patients of volunteers. Actually, among the merely 8 studies that met the first and most problematic inclusion criterion, only 5 dealt with patients, the others employed healthy volunteers, one even volunteer students, which I would argue is a population too similar to researchers to yield useful data for the authors purpose. There is no indication that this is a VERY important and relevant distinction if one wants to assess the usefulness on hand held computers for patient populations i.e., people who may have limited eyesight and dexterity, not to mention various mental challenges which all would need to be separately taken into consideration in appropriate research.

The next inclusion criterion is Compared handhelds to paper. No description of the computers used in the investigations is given, of operating systems, application software, screen sizes, image or letter sizes, data capturing software, prompts, etc, employed, training given to subjects in the studies, etc., etc. Some cursory remarks refer to Instruments and Mode of Entry but are often cryptic and almost devoid of information, such as in one of the more elaborate entries in Table 1:

Intervention: Customized version of MiniDoc, a portable electronic data capturing device with LCD (Liquid Crystal Display) screen with four push buttons.

The last inclusion criterion was

At least one of the following outcomes was assessed:
In other words, any combination of one subjective and 2 objective criteria was acceptable. The limited usefulness of subjective assessment of preferences has been well documented by Kushniruk et al. (compare, e.g., Kushniruk 2002). But even the objective criteria are dealt with in a wide range of ways in the reviewed studies. E.g., Accuracy data is reported as if it were comparable regardless of, e.g., the software and hardware employed for the portable computers, or of the basic qualification of and training provided to subjects, etc. Similar cautions apply to the objective characteristic Timeliness. One study, published in 2000, was able to shave 23% off of a data collection period of 65 minutes for each group of 10 patients. The reader is left to imagine what that might mean. Another study, published in 1991, was able to reduce the transfer of data from >30 minutes / week to 1 minutes / week. Not only do both studies obviously address time requirements for very different processes, but the fact that the two referenced publications are almost a decade apart and probably used vastly different technology remains unreported and uncommented, if not altogether unnoticed.

The paper then deals with the insights gained by these data and tries to warm the readers mind for the fact that hand held computers have the potential not only to overcome some of the limitations of conventional paper and pencil devices but also supersede them, particularly with respect to improving timeliness of data handling. In addition, the preference by research subjects for hand held computer could result in improved adherence to data collection protocols for long-term studies.

Given these statements, this reviewer has the impression that the researchers did not progress very far beyond their original assumptions. The last statement is particularly difficult to read out of the data, given that only 3 out of 8 studies assessed preferences of their subjects and that one of them dealt with Health [sic], free living volunteers (Table 1). The volunteers exhibited a weak preference for paper & pencil, the patients for the handhelds. But more to the point, there is no information on the ways in which the compared devices made people adhere to data collection protocols. There is ample indication it the more serious health information systems literature that adherence to protocols enforced by coercive software can decrease acceptance, lead to the collection of data that do not validly reflect a factual state of affairs, and may be stifling to the minds of its users. This important type of information is impossible to glean from a review of studies as heterogeneous as the ones included in the reviewed paper and illustrates splendidly that RCT design is not a helpful selection criterion. If RCT design is the selection criterion, important effects such as the ones just cited will go unnoticed. They are likely difficult to elicit even in specifically designed RCTs.

In summary: the paper is extremely weak in shedding light on the issues that the researchers pursued. One gets the impression that the authors belief in their convictions is likely fuelled by their own experimental work, which, however is only reported in an addendum to the current paper.

Essentially, the paper is a perfect illustration of the extent to which the RCT approach engenders ritualistic adherence to research routines that have little if any chance of generating insights suitable for the advancement oh health information systems. This is why this paper is important and should be published after appropriate correction of minor flaws and provided it is flanked by considerate independent reviews from researches established in the evaluation of health information systems. It is to be hoped that then the stifling practice of requiring RCT approaches for the evaluation of health information systems can be stopped. This would be very much to the benefit of the advancement of health information systems, an issue, which is currently addressed by major funding initiatives across North America and elsewhere.

Major Compulsory Revisions
The title of the paper should be:
A review of randomized controlled trials evaluating the effectiveness of hand held computers for data collection through research subjects in health care.
On page 12, third line, the beginning of the sentence currently reading In all but one of these . . . should be Two out of the three . . .
In Figure 1 font sizes may have to be adjusted so that the text in the boxes is displayed completely.
The text of entries in Table 1 is sometimes truncated or broken at random in the middle of word. Adjustments would be desirable.
The entry for Patient Population for the study by Stratton RJ et al. 1998 should probably start with the word Healthy instead of Health.
I would suggest that it is probably not justified to summarize the data in Table 4 to totals, given the vast differences in study populations and the lack of information of the processes on which preferences are expressed.

Minor Essential Revisions
The entry for Patient Population for the study by Stratton RJ et al. 1998 should probably start with the word Healthy instead of Health.

Discretionary Revisions
The authors might want to include their own recent study currently reported in an addendum in the paper proper.

Which journal?: Not appropriate for BMC Medicine: an article whose findings are important to those with closely related interests and more suited to BMC Medical Research Methodology.

What next?: Offer publication in BMC Medical Research Methodology after minor essential revisions.

Quality of written English: Acceptable.

Statistical review: No

Declaration of competing interests:
I declare that I have no competing interests.