

Model Formulation ■

Methodologic Issues in Health Informatics Trials: The Complexities of Complex Interventions

IVAN SHCHERBATYKH, MD, MPH, ANNE HOLBROOK, MD, PHARM.D, MSc, LEHANA THABANE, PhD,
LISA DOLOVICH, PHARM.D, MSc, FOR COMPETE III INVESTIGATORS

Abstract Objective: All electronic health (e-health) interventions require validation as health information technologies, ideally in randomized controlled trial settings. However, as with other types of complex interventions involving various active components and multiple targets, health informatics trials often experience problems of design, methodology, or analysis that can influence the results and acceptance of the research. Our objective was to review selected key methodologic issues in conducting and reporting randomized controlled trials in health informatics, provide examples from a recent study, and present practical recommendations.

Design: For illustration, we use the COMPETE III study, a large randomized controlled clinical trial investigating the impact of a shared decision-support system on the quality of vascular disease management in Ontario, Canada.

Results: We describe a set of methodologic, logistic, and statistical issues that should be considered when planning and implementing trials of complex e-health interventions, and provide practical recommendations for health informatics trialists.

Conclusions: Our recommendations emphasize validity and pragmatic considerations and would be useful for health informaticians conducting or evaluating e-health studies.

■ *J Am Med Inform Assoc.* 2008;15:575–580. DOI 10.1197/jamia.M2518.

Introduction

Electronic medical records (EMRs), computerized decision support systems (CDSS), electronic prescribing and other health information systems are increasingly promoted as effective tools to improve the quality and efficiency of

Affiliations of the authors: Centre for Evaluation of Medicines, St Joseph's Healthcare Hamilton (IS, AH, LT, LD), Hamilton, ON, Canada; Department of Clinical Epidemiology and Biostatistics, McMaster University (IS, AH, LT), Hamilton, ON, Canada; Division of Clinical Pharmacology and Therapeutics, Department of Medicine, McMaster University (AH), Hamilton, ON, Canada; Department of Family Medicine, McMaster University (LD), Hamilton, ON, Canada.

The COMPETE III study was funded by the Ontario Ministry of Health and Long-Term Care Primary Health Care Transition Fund #G03-02820. Dr. Ivan Shcherbatykh is supported in part by the Canadian Institutes for Health Research (CIHR) Health Informatics PhD/Postdoc Strategic Training Program and by the Father Sean O'Sullivan Research Centre, St. Joseph's Healthcare Hamilton, Canada. Dr. Anne Holbrook was supported in part by a CIHR Investigator award.

This paper was presented in part at the Critical Issues in eHealth Research Conference: Toward Quality Patient Centered Care, September 11–12, 2006, Bethesda, MD US, and the International Conference addressing Information Technology and Communications in Health 2007, February 16–18, 2007, Victoria, BC Canada.

No conflicts of interest.

Correspondence: Ivan Shcherbatykh, Centre for Evaluation of Medicines, 105 Main Street East, Level P1, Hamilton, Ontario L8N1G6 Canada; e-mail: <shcheriy@mcmaster.ca>.

Received for review: 05/22/07; accepted for publication: 04/27/08.

patient care. However, before these expensive systems are widely adopted, it is essential that they meet the usual standards of evidence of benefit versus harm, and of cost-effectiveness.¹ Of various study designs available,^{2–4} including qualitative methods, in-house testing, observational studies, and controlled trials, a well-executed prospective randomized controlled trial (RCT) is considered the strongest design for its internal validity—the unbiased assessment of cause and effect.^{5,6} This method, particularly practical RCTs, is sufficiently robust to assess benefits, harms and cost of health informatics (HI) interventions and to address the key evaluation domains of validity (“Did it really work?”), generalizability (“Can it work elsewhere?”), and cost-effectiveness (“Is it worth the cost?”). However, prospective RCTs are challenging to implement in the field of health informatics for several reasons.^{2,4,6} First, most HI trials are studies with ‘complex interventions’, meaning multifaceted, involving multiple targets (patients, clinicians) and various active components that “may act both independently and interdependently.”⁷ Figure 1 illustrates a hypothetical complex HI intervention that includes different components and factors, some of which may facilitate delivery of the intervention (e.g., ownership of CDSS, user satisfaction), while others may become a barrier to its implementation (e.g., poor integration into workflow).⁵ Rigorously designed HI interventions have to prospectively identify or account for many of these technology-related factors, which adds to their complexity and expensiveness (even compared to similarly designed implementation research or quality improvement interventions),⁸ and makes

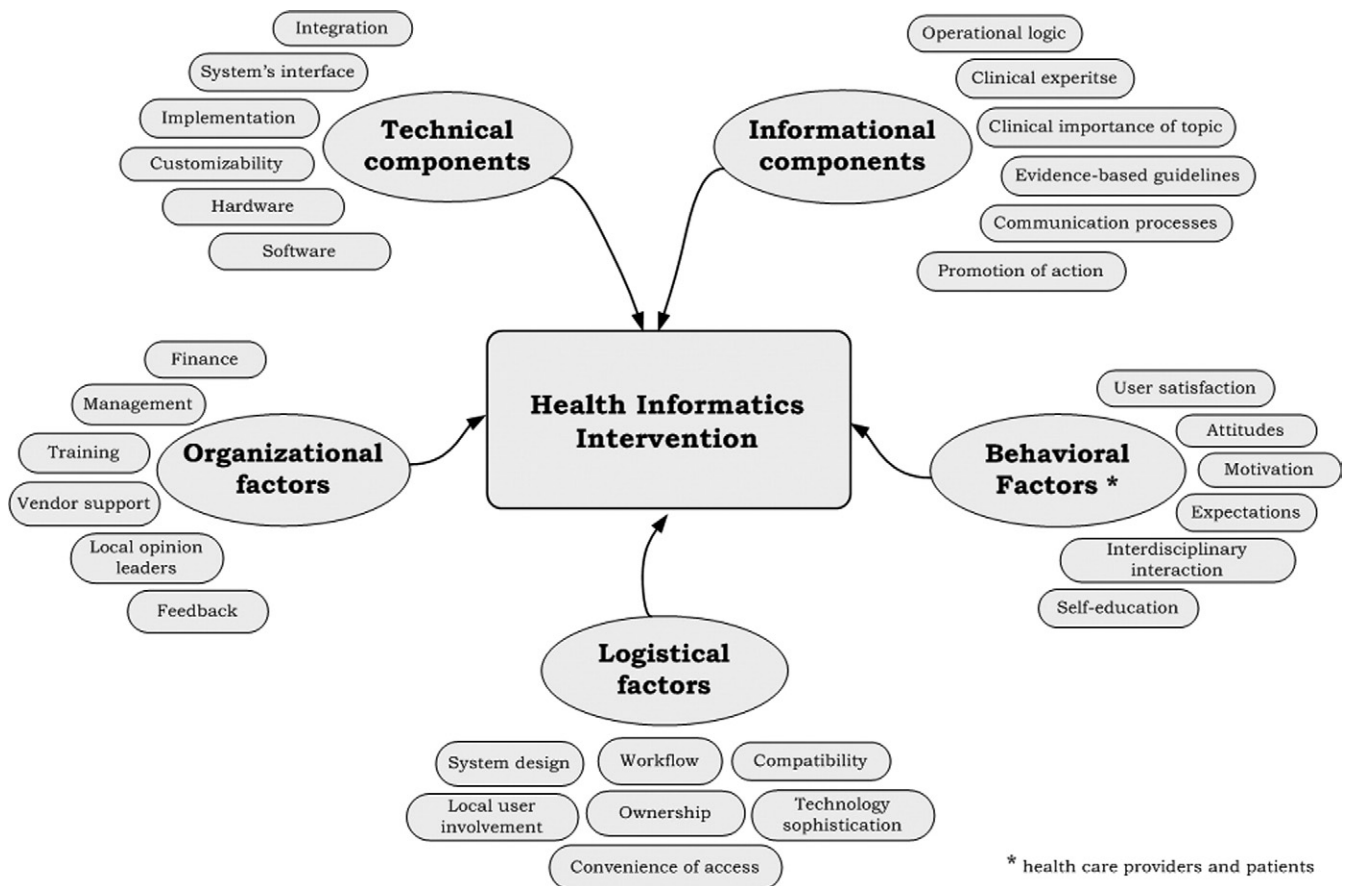


Figure 1. Components of a Complex Health Informatics Intervention.

these studies more difficult to organize, monitor and complete. Second, many of the recommended features of RCTs of therapies are not always feasible in studies of electronic health (e-health) technologies. For example, blinding of trial participants, adequate sample size, and complete follow-up are key to validity,^{6,9} but can be difficult to implement in trials of health information technology due to complexity of their interventions and “real-time” practice-based nature. Finally, comprehensive methodologic guidelines for conducting multifaceted HI studies do not exist. Health informatics also lacks specific reporting criteria, such as modifications of CONSORT or QUORUM statements for randomized trials and systematic reviews,^{10,11} which journals can request and enforce.

The objectives of this paper are to describe selected key methodologic issues in conducting and reporting randomized controlled trials in health informatics, provide real-life examples from a recent study, and present recommendations for HI trialists (Table 1).

COMPETE III Trial

Methodologic considerations described here were addressed throughout different stages of the Computerization of Medical Practices for the Enhancement of Therapeutic Effectiveness (COMPETE) III study—a pragmatic, multi-centered randomized controlled trial, which was launched in January 2005, with follow-up being completed by July 2006. Trial data are now being analyzed.

Briefly, the study investigated the impact of an individualized, patient-physician shared, web-based electronic decision support on vascular risk management integrated with practice EMRs. The trial recruited 49 family physicians in Ontario, Canada. Over 1100 patients with vascular risk factors or vascular disease (cardiac events, stroke, peripheral vascular disease, diabetes) were enrolled and randomized to participate in intervention (electronic vascular risk management program) or control arm (usual care in family practice) for a period of at least 12 months. The study hypothesized that patients in the intervention arm would lower their vascular risk more than those in usual care, reflected by a change of at least 20% in the study primary endpoint. The COMPETE III vascular management program included individualized, evidence-based monitoring and advice regarding 15 vascular prevention variables (family physician visits, blood pressure, low-density lipoprotein cholesterol, glycosylated haemoglobin HbA1c, waist:hip ratio, urine albumin, smoking, diet, exercise, psychosocial state, aspirin or antiplatelet medication use, eye exam, foot exam, influenza immunization, and medication adherence) with the vascular tracker, as well as support from clinical care coordinators (CCCs). The study was approved by the St. Joseph's Healthcare Hamilton Institutional Review Board, and all patients and physicians enrolled in the trial signed an informed consent.

Table 1 ■ Methodologic Issues and Recommendations for Health Informatics Trialists

| Methodologic Issue | Recommendations |
|---|---|
| 1. Intervention design | Use a pragmatic randomized controlled trial design if the study aims to determine effectiveness of a health informatics intervention. The following factors will determine the choice of specific design, such as two-group parallel, cross-over, 'early versus delayed', factorial, or cluster-randomized trial: 1) research question (HI intervention versus control, one intervention versus another), 2) type of intervention (CDSS, computer-generated reminder, computerized physician order entry), 3) outcome measures (patient-related endpoints, provider performance, economic analysis), 4) participants, 5) setting (primary care, tertiary care, university-affiliated), 6) length of a follow-up period (short-term versus long-term). |
| 2. Choice of randomization | Select individual-level versus cluster-level randomization considering the following factors: potential and magnitude of contamination, unit of analysis in the study (e.g., patients, physicians, hospital wards), feasibility (e.g., ability to recruit a sample size large enough to adjust for clustering, cost, ethical considerations), and existing workflow (e.g., availability of electronic prescribing in one unit versus the whole hospital). Sample size calculations and statistical analysis should be adjusted for clustering effect; also, special care should be taken to prevent possible selection bias with cluster randomization. |
| 3. Allocation concealment | Allocation concealment at the time of randomization should be done by using adequate methods (e.g., by an individual not otherwise involved in the trial). This maneuver is critical for ensuring the validity of study results, and is always feasible. |
| 4. Blinding of subjects | Determine whether designs such as 'early versus delayed' or factorial can be utilized to blind study participants. 'Partial blinding' can also be used. If blinding of participants is not possible then incorporate blinding of outcome assessors, data collectors, statisticians, or other strategies to ensure comparable experience between study groups. Use of information technologies can be used to assist with blinding. |
| 5. Components of a complex intervention | Identify and describe the active components within your HI intervention (e.g., computerized physician decision support, patient alerts, provider education), and predict the mechanisms by which these will contribute to the overall success of the study. Make sure all components are described in sufficient details. |
| 6. Sample size and power | Plan ahead and calculate sample size and power for your HI trial, using the results of a pilot study or past similar interventions. Generally, patient-related clinical outcomes require larger sample size or power, compared to process or composite endpoints. Accounting for clustering will lead to increase in sample size. Allow adjustments of sample size for protocol violations, participants drop out, and design considerations. |
| 7. Outcome measures | Choose outcome measures that are clinically relevant, sensitive, and measurable. Consider the trial duration and outcome prevalence in your sample population. Estimate or calculate minimal clinical important difference for each primary outcome. Consider validation of newly developed outcomes or scores. Evaluation of potential harm or negative effects, as well as economic analysis of health informatics intervention should be among the measured endpoints. |
| 8. Statistical analysis | Primary statistical analysis should always follow the intention-to-treat principle, preserving the power of randomization. Sensitivity analysis ("per protocol" or "on treatment") can be used as a secondary approach, especially in explanatory HI studies. Statistical analysis should reflect the study design (e.g., factorial, cluster-randomized). |
| 9. Follow-up and missing data | Ensure complete patient follow-up, including situations when they participate through virtual mechanisms. Unless this violates consent or privacy agreement, participants who drop out of study should be followed up as well (this could be negotiated in advance). There are many ways to prevent or minimize missing data in study, and investigators have to make a pragmatic decision about which strategy to choose. Missing values should be appropriately reported and handled, preferably using multiple imputation techniques. |
| 10. Reporting | Use the CONSORT guidelines to report the results of HI intervention and overall trial logistics to allow readers to interpret and generalize the trial findings to different settings, systems, or populations and to help them better understand the study's strengths and limitations. |

Methodologic Issues

The methodologic issues described here are not exhaustive nor exclusive to HI interventions; rather, based on our experience and that of other authors,^{2,5,9} we believe these issues are often overlooked in studies of health information technology. Providing this review of the topics as they apply to an e-health example illustrates how methodologic challenges can be addressed within the complexity of a HI study. Our recommendations are summarized in Table 1.

Choice of Randomization

The two types of randomization generally used in HI trials are individual-level and cluster-level. Individual-level ran-

domization is the preferred approach for most simple RCTs.¹² However, in more complex studies where 'clustering effect' may be present and where responses from individuals are not truly independent, use of individual randomization and standard statistical analysis without regard to clustering may result in false-positive results.¹³ Conversely, the HI interventions that are applied to physicians may be particularly prone to 'contamination', whereby subjects allocated to the control group are actually exposed to the intervention, or vice versa.^{12,14} Even if taken into account in the analysis, contamination may lead to dilution of the intervention effect. Cluster-level randomization (use of clusters as allocation units) can be used to counteract these

pitfalls. Since many HI interventions target physicians or hospital units, randomization at these levels seems intuitive, simplifies trial administration, and decreases contamination (if control and intervention subjects are well separated).¹³ However, numerous statistical and ethical challenges associated with cluster randomization hamper its wide use. Cluster-randomized trials have considerably reduced statistical efficiency, require inflated sample sizes and complex statistical analysis, and are difficult to meta-analyze.¹⁴ In addition, cluster RCTs can be particularly vulnerable to selection bias (systematic differences between comparison groups), which may necessitate special measures, such as prospective identification of all eligible patients prior to randomization.^{12,15}

The COMPETE III trialists, after considering both options, decided to randomize individual patients for four main reasons. First, because the intervention was complex, it was difficult to imagine that contamination would exceed 30%, the level after which contamination has an actual impact on sample size.¹² In addition, the patients themselves were the “minority target.” Third, many physicians in our study were unwilling to entertain the chance of being a control group practice. Fourth, based on previous work, we had a good estimate of intracluster correlation coefficient measuring similarity of patients within practices, which we applied to our sample size to account for potential clustering.

Blinding

Blinding, or the concealment of group assignment, helps to prevent bias in RCTs.^{11,16} Study participants that potentially could be blinded include patients, clinicians, data collectors, outcome adjudicators, statisticians, members of the data and safety monitoring committee, and manuscript authors.^{16,17} However, in many non-pharmacological studies, including HI trials, it is not possible to completely blind patients or health care providers.^{9,18} Thus, some alternative precautions must be undertaken to eliminate or reduce the chance of bias. These measures might include: allocation concealment at the time of randomization (considered as most important);¹⁹ application of co-interventions (ancillary treatments like drug therapy or counselling) to all study groups; unbiased evaluation of outcomes through blinded data collection, adjudication and analysis; use of objective primary or secondary outcomes; use of intention-to-treat (“as randomized”) analysis, and complete patient follow-up. In many pragmatic HI studies ‘partial blinding’ may be used, such as blinding of subjects to study’s hypothesis or specific outcomes.^{18,20} Blinding of participants can also be achieved with an ‘early versus delayed’ design, which allocates subjects to participate in the intervention either immediately or after some time, so subjects are unaware of their status during the timeframe being analyzed. Blinding of clinicians to the e-health intervention can occasionally be achieved: one study described blinding as programming the system not to display the alerts on control arm patients.²¹ Finally, we have observed that the precise nature and intent of a complex HI intervention is frequently forgotten by participants over time due to the intensity and pace of clinical practice, which by itself is an assist to blinding.

In COMPETE III the nature of the intervention precluded blinding of patients, family physicians, and the CCCs who

assisted the intervention patients. Therefore, steps were undertaken to reduce potential bias and ensure comparable experiences between control and intervention participants. The computer-generated allocation schedule was concealed and was accessible only by the authorized interviewers. The data collection and entry procedures followed the same protocol for both study groups. Many study outcomes were objective endpoints, measured by independent clinical laboratories, while the assessment of other endpoints was done by third party interviewers and chart auditors, not otherwise involved in the trial and uninformed of the study hypotheses. The trial investigators, statisticians, manuscript writers, and members of the Scientific Advisory Board were blinded throughout the duration of study, data analysis, and manuscript preparation.

Outcome Measures and Sample Size

Clinical, Process, and Composite Endpoints

Selection of sensitive and relevant outcome measures is challenging in health informatics studies.^{5,22} While improvement in clinical outcomes (measures of morbidity or mortality) represents the ultimate goal of all clinical interventions, important changes in these endpoints are hard to achieve in the short term, which is a usual timeframe for many HI trials. Process outcomes or ‘performance measures’ reflect what was done for the patients, e.g., investigations, treatments, or counselling.²³ These endpoints are more sensitive to change, often easier to measure and collect (especially through an EMR), and require smaller sample sizes in clinical trials.²³ Yet, process outcomes are surrogates and require validation through establishing their properties and degree of association with clinical events. Composite outcomes (combinations of several endpoints) and composite scores (combinations of several endpoints that are weighted or scored) are popular in HI interventions as they combine different manifestations of chronic or complex diseases, and enable the use of smaller sample sizes or shorter follow-up periods.^{24,25} However, composite outcomes and scores are often hard to interpret and validate.²⁴ In addition, investigators have to ensure that individual components within a composite outcome are clinically relevant endpoints, with comparable rates of occurrence and similar importance to patients.²⁵ Determining the minimal clinically important difference (MCID) for a composite outcome as required in sample size calculations is also challenging, as the rates of occurrence and clinical significance often are unknown. Limited recommendations for composite measures construction and validation presently exist in HI trials,^{25,26} and their reporting in trials is often poor or misleading.²⁵

The COMPETE III used a process composite score (PCS) to compare the intervention and control groups. The PCS was the sum of 10 vascular tracker process outcomes (blood pressure, cholesterol, glycosylated haemoglobin HbA1C, waist:hip ratio, urine albumin, smoking, diet, exercise, psychosocial state, aspirin or antiplatelet medication use) that are most strongly supported by the evidence in management of vascular and diabetic risk patients.^{27,28} Each process outcome was assigned a score of 1 and weighted depending on recommended frequency of monitoring for the follow-up period of 12 months. We used a combination of critical appraisal and consensus amongst investigators to establish

minimal clinically important changes for the PCS components. This approach is one of the established methods for MCID ascertainment,²⁹ and was used to determine initial face and content validity of the PCS components. Multiple secondary outcomes are common in complex intervention trials. The COMPETE III secondary outcomes include clinical outcomes, composite components, patient goals, medication adherence, and facilitators and barriers to implementation of the intervention. A cost-effectiveness and willingness-to-pay analysis are underway to determine if the increased cost of the intervention is reasonable compared to health gains.

Sample Size

An adequate sample size is essential for validity of clinical studies, but is not easily obtainable in health informatics trials.^{4,5} Common reasons include lack of access to the technology, low computer literacy among the participants, and complicated informed consent procedures.³⁰ Achieving sufficient sample size may also be difficult when HI trialists randomize physicians or health care teams instead of individual patients.⁹ Similar to other studies, sample size and power calculation in multifaceted HI interventions is further complicated by adjustments for clustering, use of multiple primary outcomes, protocol violations, and loss of participants.

The COMPETE III staff used multiple channels (e.g., EMR vendors, governmental listings, media advertising) to recruit family physicians into the study. Patients were identified through physician's EMR and billing systems; all eligible patients regardless of their computer skills were invited to the trial. Since there were no published studies with a similar primary outcome, sample size assumptions in COMPETE III were based on the data from our previous projects. We estimated the treatment effect as *minimal difference worth detecting*—the minimum difference that would lead to change in management of patients with cardiovascular disease. The COMPETE III trial aimed to enroll at least 1100 patients in total, allowing detecting 20% difference in the process composite score between the two groups. This sample size was based on a conservative estimate of intra-cluster correlation coefficient to account for potential clustering among the physicians.

Adverse Effects or Harm

Outcome measures usually focus on potential benefits, but should be adequate to measure and have power to reliably detect at least the important adverse effects. Use of CDSS may increase a physician's consultation time or otherwise interfere with the routine practice activities, or even provide misleading advice. Computerized drug interaction alerts are well known to be suboptimal by producing too many false positive warnings.³¹ The HI trials do need to collect feedback on the negative impact of the intervention as well as the positive—few reports in the past outlined the various ways that poorly developed or implemented CDSS and physician order entry systems could harm patients.^{32,33} Yet these negative aspects must be identified and analyzed to drive improvements. General guidelines for reporting harm, such as extension of the CONSORT statement, are also recommended.³⁴

Our experience has shown that questionnaires directed at ease of use and perceived usefulness of the technology, satisfaction with the study, willingness to pay for the intervention, combined with qualitative studies on facilitators and barriers for generalized use, are useful to pick up adverse effects. More challenging with distributed community sites is the prompt identification and repair of technical problems, which can rapidly erode the use of the intervention but may not be spontaneously reported for some time. We recommend future work on a comprehensive, standardized set of outcomes and indicators that would be used for measurement and evaluation of harm in e-health studies.

Follow-up and Missing Data

Incomplete or missing data are a constant challenge in clinical studies. The HI trials might be more prone to missing data due to the settings being routine clinical practice, multiple data sources (paper charts, web-based forms, fax forms, EMR data, billing software), complexity of interventions, and multiple trial participants. Large amount of missing values makes it difficult to draw reliable inferences and may bias the results;¹¹ incomplete data in CDSS may result in inappropriate or unsafe recommendations.³⁵ Potential solutions for HI trialists may include attempts to minimize missing data at the design stage (e.g., negotiating with patients to allow continuous data collection in case of their withdrawal), utilization of intention-to-treat analysis, and use of well-accepted methods of missing data handling, such as multiple imputation. This statistical technique for analyzing incomplete data has been shown to be a superior approach to simple imputation methods in missing data analysis.³⁶

The COMPETE III trial used several strategies to minimize incomplete data and to handle missing data appropriately. First, we tried to reduce any loss of information by paying attention to patient and physician compliance. This was done by using behavioral approaches through incentives and feedback. The CCCs were monitoring the extent to which the patients kept regular appointments with their family physicians and re-filled prescriptions, while the telephone reminder system provided prompts for the patients. We also standardized the data collection and provided baseline training for the chart reviewers. The study followed the intention-to-treat principle and accounted for all randomized participants. Finally, we had a number of security measures to prevent any unauthorized access to data or data loss (e.g., safes for data storage, on-line data security). Given that missing data are inevitable despite these efforts, we have developed standard multiple imputation routines.

Discussion and Recommendations

In this paper, we reviewed selected key methodologic issues that are often overlooked in practical RCTs in health informatics,^{4,5,10,30} and how they were handled in COMPETE III study. Our e-health team considered these challenges as the most important. We realize that the list of issues is not fully comprehensive; yet, we consider that proper implementation and reporting of these features should be a 'minimum requirement' for every rigorously designed HI trial. Our recommendations summarized in Table 1 expand upon prior guidelines for RCTs of therapies and complex interventions,^{3,6,7,10} and would be useful for health informaticians

conducting or evaluating HI studies since they emphasize both validity and pragmatic considerations, and provide e-health examples. Research in complex health informatics interventions could also benefit from a CONSORT-like statement that established guidelines for conducting and reporting of these types of trials.

We continue to recommend that all health information technologies (electronic medical records, electronic prescribing, sophisticated decision support systems) should be required to demonstrate in an unbiased evaluation that they provide more benefit than harm on patient-important outcomes and are likely to be worth the cost once fully developed and deployed. The latter point is especially important given the billions of dollars to be spent on e-health technologies, money that could be wasted if early, high quality evaluation of clinical impact is not carried out.

References ■

- Eisenstein EL, Ortiz M, Anstrom KJ, Crosslin DR, Lobach DF. Economic evaluation in medical information technology: why the numbers don't add up. *AMIA Annu Symp Proc*. 2006;914.
- Rosenbloom ST. Approaches to evaluating electronic prescribing. *J Am Med Inform Assoc*. 2006 Jul-Aug;13(4):399–401.
- Eccles M, Grimshaw J, Campbell M, Ramsay C. Research designs for studies evaluating the effectiveness of change and improvement strategies. *Qual Saf Health Care* 2003 Feb;12(1):47–52.
- Friedman CP, Wyatt JC. *Evaluation Methods in Biomedical Informatics*. 2nd ed. New York: Springer-Verlag; 2006.
- Garg AX, Adhikari NK, McDonald H, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *JAMA* 2005 Mar 9;293(10):1223–38.
- Guyatt GH, Haynes RB, Jaeschke R, et al. The philosophy of evidence-based medicine. In: Guyatt GH, Rennie D, editors. *Users' guides to the medical literature: a manual for evidence-based clinical practice*. Chicago, IL: AMA Press; 2001. p. 3–12.
- Campbell M, Fitzpatrick R, Haines A, et al. Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000 Sep 16;321(7262):694–6.
- Foy R, Eccles M, Grimshaw J. Why does primary care need more implementation research? *Fam Pract*. 2001 Aug;18(4):353–5.
- Randolph AG, Haynes RB, Wyatt JC, Cook DJ, Guyatt GH. Therapy and validity: computer decision support systems. In: Guyatt GH, Rennie D, editors. *Users' guides to the medical literature: a manual for evidence-based clinical practice*. Chicago, IL: AMA Press; 2001. p. 291–308.
- Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med*. 2001 Apr 17;134(8):663–94.
- Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomized controlled trials: the QUOROM statement. *Lancet*. 1999 Nov 27;354(9193):1896–900.
- Torgerson DJ. Contamination in trials: is cluster randomisation the answer? *BMJ* 2001 Feb 10;322(7282):355–7.
- Donner A, Klar N. *Design and analysis of cluster randomization trials in health research*. London: Arnold; 2000.
- Chuang JH, Hripcsak G, Heitjan DF. Design and analysis of controlled trials in naturally clustered environments: implications for medical informatics. *J Am Med Inform Assoc*. 2002 May-Jun;9(3):230–8.
- Puffer S, Torgerson D, Watson J. Evidence for risk of bias in cluster randomized trials: review of recent trials published in three general medical journals. *BMJ* 2003 Oct 4;327(7418):785–9.
- Devereaux PJ, Bhandari M, Montori VM, Manns BJ, Ghall WA, Guyatt GH. Double blind, you have been voted off the island! *Evid Based Ment Health* 2002 May;5(2):36–7.
- Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. *Lancet* 2002 Feb 23;359(9307):696–700.
- Boutron I, Guittet L, Estellat C, Moher D, Hrobjartsson A, Ravaud P. Reporting methods of blinding in randomized trials assessing nonpharmacological treatments. *PLoS Med*. 2007 Feb 20;4(2):e61.
- Altman DG, Schulz KF. Statistics notes: Concealing treatment allocation in randomized trials. *BMJ* 2001 Aug 25;323(7310):446–7.
- Tamblyn R, Huang A, Perreault R, et al. The medical office of the 21st century (MOXXI): effectiveness of computerized decision-making support in reducing inappropriate prescribing in primary care. *CMAJ* 2003 Sep 16;169(6):549–56.
- McGregor JC, Weekes E, Forrest GN, et al. Impact of a computerized clinical decision support system on reducing inappropriate antimicrobial use: a randomized controlled trial. *J Am Med Inform Assoc*. 2006 Jul-Aug;13(4):378–84.
- Delpierre C, Cuzin L, Fillaux J, Alvarez M, Massip P, Lang T. A systematic review of computer-based patient record systems and quality of care: more randomized clinical trials or a broader approach? *Int J Qual Health Care*. 2004;16(5):407–16.
- Eddy DM. Performance measurement: problems and solutions. *Health Aff (Millwood)*. 1998 Jul-Aug;17(4):7–25.
- Freemantle N, Calvert M, Wood J, Eastaugh J, Griffin C. Composite outcomes in randomized trials: greater precision but with greater uncertainty? *JAMA* 2003 May 21;289(19):2554–9.
- Montori VM, Permyer-Miralda G, Ferreira-Gonzalez I, et al. Validity of composite end points in clinical trials. *BMJ* 2005 Mar 12;330(7491):594–6.
- Ferreira-González I, Busse JW, Heels-Ansdell D, et al. Problems with use of composite end points in cardiovascular trials: systematic review of randomised controlled trials. *BMJ* 2007 Apr 14;334(7597):786.
- Majumdar SR, Johnson JA, Bowker SL, et al. A Canadian consensus for the standardized evaluation of quality improvement interventions in Type-2 diabetes: Development of a quality indicator set. *Can J Diabetes*. 2005;29:220–9.
- Yusuf S, Hawken S, Ounpuu S, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet*. 2004 Sep 11–17;364(9438):937–52.
- Beaton DE, Boers M, Wells GA. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. *Curr Opin Rheumatol* 2002 Mar; 14(2):109–14.
- Dansky KH, Thompson D, Sanner T. A framework for evaluating eHealth research. *Eval Program Plann*. 2006 Nov;29(4):397–404.
- Shah NR, Seger AC, Seger DL, et al. Improving acceptance of computerized prescribing alerts in ambulatory care. *J Am Med Inform Assoc*. 2006 Jan-Feb;13(1):5–11.
- Koppel R, Metlay JP, Cohen A, et al. Role of computerized physician order entry systems in facilitating medication errors. *JAMA* 2005 Mar 9;293(10):1197–203.
- Coiera E, Westbrook J, Wyatt J. The safety and quality of decision support systems. *Methods Inf Med*. 2006;45 Suppl 1: 20–5.
- Ioannidis JPA, Evans SJW, Gotzsche PC, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004;141:781–8.
- Berner ES, Kasiraman RK, Yu F, Ray MN, Houston TK. Data quality in the outpatient setting: impact on clinical decision support systems. *AMIA Annu Symp Proc*. 2005;41–5.
- Shrive FM, Stuart H, Quan H, Ghali WA. Dealing with missing data in a multi-question depression scale: a comparison of imputation methods. *BMC Med Res Methodol*. 2006 Dec 13;6:57.