Ethical Challenges in Implementation Research

Ruth Macklin*, Department of Epidemiology & Population Health, Albert Einstein College of Medicine

*Corresponding author: Ruth Macklin, Department of Epidemiology & Population Health, Albert Einstein College of Medicine, Bronx, NY 10461, USA. Tel.: +01 718 430 3574; Email: ruth.macklin@einstein.yu.edu.

Implementation research is increasingly common in developing countries as a way of studying the introduction to the population of health interventions that have been proven to be effective elsewhere. Implementation studies are often conducted as cluster randomized trials, a design that raises ethical and conceptual questions different from those in conventional randomized controlled trials. It is often unclear who the subjects of the research are, informed consent may be difficult or impossible to obtain and controversy surrounds the use of comparison clusters that provide substandard care to the population where the research is carried out. An examination of protocols for this type of research reveals uncertainty on the part of researchers themselves about whom or what they are studying and from whom (if anyone) informed consent is required.

A type of public health research that has received scant attention in the literature of research ethics is *implementation research* (IR). Although IR can be conducted anywhere, it is especially prevalent in low- and middle-income countries (LMICs), where numerous efforts seek to improve existing public health programs. A definition of IR appears in an article that distinguishes several types of public health research often confused with one another:

> Implementation research aims to develop strategies for available or new health interventions in order to improve access to, and the use of, these interventions by the populations in need . . . [T]he starting point is the availability of an intervention or intervention package that has been proven efficacious in previous research, but for which major questions remain as to how to scale up the intervention and ensure effective integration within the health system (Remme, 2010: 4).

The World Health Organization (WHO) has recognized the importance of IR in establishing an IR platform. According to its Web site, the platform supports research that:

- Identifies common implementation problems and their main determinants that hinder effective access to interventions,
- Develops and tests practical solutions to these problems that are either specific to particular health systems and environments or that address a problem common to several countries in a region and
- Determines the best way of introducing these practical solutions into the health system and facilitates their full-scale implementation, evaluation and modification as required (World Health Organization, 2013).

Following the introduction of the IR platform, WHO issued a Practical Guide describing the important contributions IR can make to maximize the beneficial impact of health interventions in LMICs (Peters *et al.*, 2013). The Guide not only describes the various forms that IR can take, but it also says why this form of research is needed, who should be involved, what approaches and methods are appropriate and how the potential of IR can be realized. The Guide notes that IR 'can be of enormous value to a range of stakeholders from ministerial-level decision-makers, who may use implementation research to inform health policy formation, to program managers seeking to understand context-specific issues, and health providers looking to assess performance, make changes or introduce innovations' (Peters *et al.*, 2013: 19). However, neither the IR platform nor the Practical Guide addresses an array of ethical issues that can arise in the design and conduct of IR.

doi:10.1093/phe/phu003
Advance Access publication on 24 February 2014

Some ethical issues in IR are related to methodological features, whereas others pertain to informed consent. Does the methodology involve a comparison between current practice and an intervention that has been proven elsewhere, and if so, what are the ethical implications regarding equipoise? Is a randomized design necessary to achieve the best results, or may the use of historical controls provide adequate answers to the research question? Does the research propose to study the behavior of healthcare personnel who are being trained in a new intervention and if so, is their informed consent required? IR studies can involve randomization, sometimes in the form of clusters, yet other designs do not involve randomization. Such variations prompt the question whether randomization constitutes a criterion for determining that the activity is research. This is important to distinguish IR from quality improvement (QI) or program evaluation, as the latter activities are not typically subject to review by a research ethics committee (REC), whereas prior review by an independent properly qualified committee is an almost universal requirement for research involving human participants. IR differs in various ways from more traditional research, so confusion or uncertainty may arise regarding these and related questions. As the WHO Practical Guide acknowledges, IR 'presents some obvious taxonomic challenges . . . It is probably not surprising then that there is some confusion regarding nomenclature as well as significant debate regarding the scope of implementation research' (Peters *et al.*, 2013: 27).

Using several real and hypothetical examples, this article illustrates some confusion and uncertainty that may confront researchers and planners involved in IR, as well as RECs that review the proposed research. The analysis draws some tentative conclusions while recognizing that gray areas exist and controversies may not readily be resolved where uncertainty persists.

Equipoise in IR

A subset of IR studies is designed as cluster randomized trials (CRTs), a topic that has received attention in recent literature (Sim and Dawson, 2012; Weijer *et al.*, 2011, 2012). In contrast to the randomization of individuals in a randomized clinical trial (RCT), 'a cluster-randomized trial randomizes at the social group level (e.g. village, hospital, school)' (Sim and Dawson, 2012: 480). Several authors concluded a series of articles devoted to ethical issues in CRTs with recommendations in a document entitled 'The Ottawa Statement on the Ethical Design and Conduct of Cluster Randomized Trials' (hereinafter, Ottawa Statement) (Weijer *et al.*, 2012). The Ottawa Statement includes the following recommendation about control groups in CRTs: 'Researchers must adequately justify the choice of the control condition. When the control arm is usual practice or no treatment, individuals in the control arm must not be deprived of effective care or programs to which they would have access, were there no trial' (Weijer *et al.*, 2012: 7). In IR, however, whether a study uses a cluster design or randomizes individuals, subjects in the control arm may not have access outside the trial to an intervention that has been proven to be beneficial elsewhere. This is the old placebo problem in a new guise. When the control arm receives 'usual care' that is substandard, in cases such as poor infection control practices, it would be unethical to withhold the proven intervention from the controls. In these cases, a possible strategy is to use historical controls, despite the widespread view among methodologists that the methodology is significantly inferior to the gold-standard RCT. In cluster trials, all facilities or communities would receive the method proven elsewhere to be effective, and the study must be carefully designed to ensure to the extent feasible that the historical controls are similar in as many respects as possible to the intervention group.

The Ottawa Statement and the articles leading up to it do not make a clear distinction between CRTs that are IR and those that study an experimental intervention. The authors use the terms 'knowledge translation' and 'quality improvement' in referring to one common type of IR. They write:

> It is true that knowledge translation and quality improvement studies seek to improve patient care. But the fact that an educational or quality improvement intervention is being evaluated in a CRT suggests that its effectiveness is unproven. Indeed, if it was known at the start of the trial that the study intervention is effective, the CRT would be unethical' (McRae *et al.*, 2011a: 11).

This conclusion places all CRTs in the category of experimental interventions in one arm of the clusters. By considering interventions in IR 'unproven', the authors seek to retain the ethical requirement of clinical equipoise and can claim that the requirement is met. However, there is reason to question the Ottawa group's determination that IR involves 'unproven' interventions. On the conceptual level, their position departs from the definition of 'IR' cited earlier in this article, which explicitly states that the intervention

being implemented has been proven to be efficacious. More importantly, however, it confuses two types of research: one in which the efficacy of a new intervention is being tested, and the second in which the implementation of an intervention with proven efficacy is being studied. An IR study may be designed to compare implementation or scale-up of existing interventions, yet there may be debate or uncertainty regarding whether the existing interventions constitute a 'standard of care'—an ambiguous concept. The phrase can mean 'what is routinely—or standardly—done'; or it can have a normative meaning, that is, adherence to a standard of some sort, for example, one set by a professional body. When the intervention arm is studying implementation of a proven effective intervention and control facilities or units are observed but receive no intervention, the ethical question is the same as that in placebo-controlled RCTs. Is it ethical to withhold a proven intervention from a resource-poor community in the design of implementation CRTs?

An example is a series of placebo-controlled trials on screening for cervical cancer that was carried out in India in the mid-to-late 2000s. The purported aim of the trials was to study the method that uses visual inspection of the cervix following staining with acetic acid (VIA) to determine the efficacy of this method in a low-resource setting. However, the actual purpose was to train healthcare personnel in how properly to use this screening method. According to an article that criticized the ethics of the study:

> The researchers in these trials have argued that only a "no care" control arm can give definitive results and this information is essential to guide policies and programmes . . . VIA has been researched at least since the early 1990s. VIA is an affordable screening test, and there is evidence suggesting that it works about as well as the Pap smear' (Srinivasan, 2013: 149).

The flaw in the researchers' defense is that the efficacy of VIA was already well established. According to a WHO consultation report in 2002, 'The test performance of VIA suggests that it has similar sensitivity to that of cervical cytology in detecting CIN, but has lower specificity. Further research is required to improve its specificity without compromising sensitivity' (WHO, 2002). The WHO report also pointed out the need for training personnel in the use of the method, as well as for developing standard procedures for quality control. Therefore, what was needed was *not* an efficacy study of VIA but rather a study of its implementation in a new setting.

Researchers in India were not studying whether VIA was an accurate screening method for cervical cancer; they were not studying efficacy. Rather, they were studying whether previously untrained healthcare workers (HCWs) could learn how to use the method properly. With regard to the VIA intervention, clinical equipoise was not satisfied in this study, which is why the placebo-control design was unethical. Arguably, however, there may have been genuine uncertainty whether the VIA test administered by newly trained HCWs was better than no intervention. Even so, the solution is not to insist, like the Ottawa statement, that the intervention arm in IR is unproven. Two possibilities are (i) to abandon the gold-standard RCT for this type of research and instead use the second-best method, historical controls, or (ii) to use the cervical cytology method in the control arm. A problem with option (ii) is that the study would have to be done in a facility that has the capability and trained personnel to use the cytology method. Policy makers in rural and underserved areas might insist that studies be conducted in the setting where the new screening method is to be introduced.

IR and QI

An intervention was designed to improve the quality of care, typically conducted in hospitals and other health facilities and is known as QI. An examination of numerous examples of IR using the CRT design reveals that for many the goal is QI. The boundary between research and QI studies can be blurred (Lynn, 2004), giving rise to questions about the need for prior review by an REC and whether informed consent is required from anyone. Therefore, a threshold question is whether at least some cases of QI should be categorized as human subject research, or whether some or all of the usual requirements for research involving human participants may be forgone. IR studies that involve randomization meet one widely accepted criterion for an activity to count as research. But should an IR project studying the same thing without randomization be treated as research or as QI?

Consider the following two situations.

Situation 1

Health authorities in a rural health district in a developing country are concerned about the rate of infection in women who deliver in birthing centers. They design a CRT to study the results of training for health workers in the centers. Six clinics in one province are randomized for researchers to look at outcome data. Three clinics

will be provided with written materials for the staff that describe best practices in the use and cleaning of instruments and care for women during and after delivery. The clinics will be visited by an obstetrician and obstetrical nurse who will demonstrate the best practices. The other three clinics will not receive the intervention. Data will be collected from patients' records in both clinics after 6 months to compare rates of infection and other relevant outcomes. The health department wants to know whether it is worth spending money to introduce the intervention in all birthing centers in the province.

Situation 2

In a rural area in another part of the country, the health authorities decide to implement the same intervention in all existing birthing centers. The official in charge contends that it would be unethical to withhold training in infection control from any of the birthing centers because prevention of infections in women giving birth in the centers is an urgent need. The plan is to compare the rate of infections over a 6-month period with the rate of infections in the same period the previous year.

Situation 2 is an example of a study typically considered to be QI rather than research. This is because no randomization is involved, and health officials have the authority to introduce new programs they believe to be essential for ensuring or improving public health. However, as noted above, no bright line exists between QI and research, and it is reasonable to consider situation 2 as a 'blend' of the two, namely, 'QI research'. Although QI normally does not involve review by an REC, some institutions may require ethics review either by an REC or another body created for this purpose. Such review could involve examination of the qualifications of those brought in to provide education and training of the health workers, as well as ensuring protection of confidentiality of the data collected from patients' records.

Whether informed consent should be obtained from the HCWs who undergo the training is a separate question, one that might be answered differently if the activity is considered QI or research. Here again, health officials have the authority to institute new public health programs without obtaining informed consent from the workers who are assigned to implement those programs. Even under the 'blended' category of QI research, a waiver of consent can be justified. The 'research' component is the comparison of rates of infection in women giving birth during the intervention and previously. The actual behavior of the HCWs is not being observed or recorded in the QI study described in situation 2.

However, implementation CRTs like in situation 1 do involve randomization, a feature that serves as a clear criterion for research. IR of this type poses two questions that do not (or only rarely) arise in the traditional research context: Who are the subjects? And when is obtaining informed consent unfeasible or likely to bias the results?

In one article, the authors of the Ottawa Statement correctly observe that the answer to the question who is the research subject in a CRT in health research may vary, depending on the study design, population or intervention under investigation. The article defines 'research subject' as 'an individual whose interests are put at risk as a result of interventions in a research study' (McRae *et al*, 2011b: 5). While this definition makes good sense, it is also reasonable to consider as human subjects of research those individuals at whom the intervention is targeted even when they are not placed at risk. In situation 1 above, it is the (HCWs) who are targeted by the intervention. Although one normally thinks of patients as the bearers of risks in research, when the behavior of HCWs is being studied, they too may be placed at risk. Anyone found to be engaged in substandard medical or nursing care is arguably at risk for sanctions of some type, which may include dismissal or charges of malpractice. The usual safeguard promised to any subjects of research involving information that may harm them if released is protection of their confidentiality. However, if a study reveals that some HCWs are providing substandard care, preserving confidentiality conflicts with the duty of a health facility to avoid foreseeable harms to patients. In a design in which the individual behavior of HCWs is not recorded and the only data used for the research are patients' records, then the HCWs are not placed at risk but are still the ones targeted by the intervention. Therefore, both definitions of 'research subject' are potentially applicable, depending on details of the particular study.

Because the HCWs in situation 1 are the subjects of research, the usual requirement would be to obtain their informed consent and allow informed refusal to participate. But the cluster design poses a problem for the usual requirement. If some of the HCWs refused to participate, it could make the results of the study uninterpretable. Some would receive the training in best practices, but those who refuse to participate would not. Outcome data obtained from patients' records could not accurately reflect the care they received at the hands of this mixed group (trained and untrained) HCWs. How could one determine whether the training of HCWs was ineffective or whether failure to significantly lower the rate of infection was due to lack of training of some HCWs? This dilemma of informed consent is a feature

of the CRT design and does not pertain specifically to IR. One practical solution is for the REC to waive the requirement for informed consent on grounds that the research would otherwise not be feasible and that the HCWs would be subject only to minimal risks. A much less desirable alternative (perhaps not practicable) would be to seek informed consent from the HCWs and transfer to one of the control clusters any workers who are unwilling to undergo the training. A third option—questionable on conceptual grounds—is to claim that despite the randomization of clinics, this is a QI project and does not constitute human subject research. The first option is both practically optimal and ethically acceptable.

Patients themselves may be individuals 'whose interests are put at risk' in IR, depending on the intervention. However, 'being put at risk' is not a sufficiently precise criterion for determining who is a research subject in IR when the only involvement of patients is the examination of their records. No interaction with the patients occurs, and confidentiality protections should be in place. In situation 1, patients' records are reviewed in both the control clusters and the intervention clusters. Patients in the intervention cluster may be treated differently by HCWs who receive the intervention, but the intervention has previously been proven effective in reducing infection so patients are not being put at additional risk. Even if the patients are considered research subjects, consent may be waived because their records are being examined only retrospectively, and confidentiality protections should be in place to minimize any risks of disclosure. However, patients' consent would be needed if part of the study involves interviews with them (e.g. concerning how they were treated by HCWs), as the purpose of such interviews is not to learn information relevant to the patients' diagnosis or treatment. When patients are interviewed as part of research—even if they are not the 'primary' subjects—their consent should be obtained. Their consent may be oral—it need not be written—and the information provided should be limited to the purpose of the interview, its contents and the length of time the interview will take, along with the usual confidentiality protections. Patients could refuse to be interviewed, and that would not have consequences for the main data being collected in the study.

IR and Public Health Practice

Some IR falls within a gray area between public health practice (e.g. surveillance, program evaluation) and public health research. Despite the attempt by the US Centers for Disease Control and Prevention to distinguish clearly between these two types of public health activity (Centers for Disease Control and Prevention, 2010), uncertainties remain. Consider the following two situations.

Situation 3

The Ministry of Health in a developing country seeks to lower the incidence and prevalence of diarrheal disease by studying whether adding a proven disinfectant to the water supply is effective. Although articles in the public health literature provide unmistakable evidence for the effectiveness of the disinfectant, the ministry of health is uncertain whether so many other factors contribute to diarrheal disease in the region that disinfecting the water would make only a small difference. In some communities, the disinfectant is added to the water supply, while in other communities, the control clusters are left alone. The communities to receive the intervention are randomly selected. If the rate of diarrheal disease is sufficiently lower in the intervention communities, the ministry will use a portion of its limited budget to introduce the disinfectant in as many communities as possible.

Situation 4

In a neighboring country, having read the public health literature showing that adding disinfectants to the water reduces diarrheal disease, the health minister decides to distribute the proven disinfectant for the water supply in all communities with the highest rates of diarrheal disease. She plans to compare the rate of diarrheal disease several months after the intervention with the same months the previous year.

In situation 3, randomization constitutes the criterion for calling the study research. Is this human subject research? Are members of the intervention communities subjects? Are members of the control communities subjects? It is obvious that obtaining informed consent from members of either group would be impossible. Yet it is clearly IR according to the definition cited earlier, as well as the WHO criteria, as the intervention has previously been proven to be effective. What about situation 4? It would appear to be a public health implementation program, which still involves looking at cases of diarrheal disease. Does examination of specific cases make it human subject research? Or is it public health research without a human subject component? Although the absence of randomization does not by

itself render a study nonresearch, it could just as well be termed 'program evaluation.' Even when informed consent is impossible because of the nature of the intervention, as in these scenarios, a question remains about the need for prior ethics review. Public health practice normally does not undergo such review, but if an activity is determined to be research, it would require review by a duly authorized review body.

There are good reasons to consider instituting an ethics review of public health activities that are not considered research. In situation 4, for example, such review might involve ensuring that the workers responsible for adding the disinfectant to the water are properly trained, that the confidentiality of people whose records are examined for diarrheal disease is adequately protected and that resulting benefits are disseminated to other communities at similar risk.

Conceptual and Ethical Confusions

As noted in the preceding sections, two activities that blur the line between research and nonresearch are public health practice and QI (Fairchild, 2003; Lynn, 2004). The methodology may be identical in activities that are considered to be nonresearch and those typically considered research, and publication of results may follow the completion of both kinds of activity. [Of note, a pertinent Centers for Disease Control and Prevention policy says that publication of findings does not differentiate research from nonresearch (Centers for Disease Control and Prevention, 2010: 3)]. Moreover, informed consent may be needed for both types of activities, but that requirement remains uncertain in cases treated as nonresearch, and, of course, the requirement for prospective review of research by an REC would normally not apply to activities that are public health practice or QI. This gray area causes quandaries for investigators, not knowing whether they need to develop a full protocol to submit to an REC. It also causes uncertainty for REC administrators and members, prompting debates among committee members and sometimes disagreements between researchers and RECs to which they must submit research protocols. These same quandaries exist regarding IR, which has received less attention in the research ethics and public health literature. As noted earlier in discussing situations 2 and 4, public health activities that qualify as nonresearch may still pose risks to individuals or communities, and therefore are appropriate for ethics review.

An examination of a number of project proposals to conduct IR revealed that in some cases, investigators were not entirely clear what they are studying and how to formulate the research questions. The researchers themselves had to be educated by program officers in their institution. The following example is illustrative.[1]

Researchers designed a study in a low-resource environment to determine which of two existing models for treating patients with an infectious disease works better. The study design takes the individual patient as the unit of analysis, with data to be collected about patients who are lost to follow-up and their levels of adherence to the medical treatment regimen. At the same time, the study plans to randomly sample 60 health facilities where the two different models are used to look at additional variables at the different sites. The protocol describes the goal as providing greater understanding of best practices in health services delivery and informing decision making about programming.

A review of the protocol by a program officer at the sponsoring organization called for changing the research question. Instead of asking which models work better, the reviewer suggested that a more appropriate IR question for the study would be why do some models work better than others. Further, the review noted that in light of the need to reframe the research question, the unit of analysis should not be the individual patient but instead the facilities where the treatments are provided. Although it is acceptable and necessary to monitor patient outcomes, the reviewer said that should not be the primary outcome for the study. Rather, the focus should be on the service delivery infrastructure and understanding how the different models work in these contexts. Given the shift to implementation questions, the reviewer recommended reducing the number of facilities in the study.

The examination of these protocols designed as IR in LMICs revealed that in some cases the investigators inappropriately took the individual patient as the unit of analysis when they should have realized they were studying the behavior of HCWs at the facility level. In other cases, the researchers were uncertain whether they were designing a pilot study or doing IR. The WHO Practical Guide makes a critical distinction between these:

> Too often interventions that work in small-scale pilot studies fail to live up to expectations when rolled out in national strategies, or fail to transfer from one country to another as a result of contextual differences. Implementation research . . . helps to clarify why that happens . . . (Peters *et al.*, 2013: 8).

In the cases where researchers thought they were conducting a pilot study, program officers in the sponsoring

organization tried to clarify by querying them about whether the proposed intervention was unproven or whether there was existing evidence to support its implementation in a new setting. Evidently, the sponsoring organization considered it important to distinguish between IR and research involving unproven interventions. As one review stated: 'It will be good to rephrase the research so that it reflects the implementation and scale up. As of now, it reads like the testing of an intervention.'[2] In another CRT, the sponsoring agency's review noted that the proposed study design was not needed to address the implementation questions that should be asked in the study: 'The main purpose of this research is not to test a new intervention, but rather to determine how to implement the intervention.' Here again, as in other IR studies, reviewers noted that the impact measures should not be patient-level outcomes but rather outcomes related to service delivery. That determination led to the question of informed consent: from whom—if anyone—is consent needed in IR? The two criteria noted earlier—individuals targeted by the intervention and individuals placed at risk—may diverge.[3]

Some IR that targets HCWs may place patients at risk. A leading example is the introduction of a diagnostic or therapeutic instrument of proven efficacy that is being introduced for the first time in a low-resource setting. The HCWs have to be trained in the procedure, which may carry some risk to patients if not used correctly. Although the intervention is targeted at the HCWs, the patients may be at risk at the hands of the workers learning the procedure. In situations like this, both the HCWs and the patients can be considered research subjects, as both definitions are applicable.

Conclusion

The reason why an understanding of IR is important has been summarized succinctly by the WHO: 'A key challenge faced by the global health community is how to take proven interventions and implement them in the real world' (Peters *et al.*, 2013: 8). The contributions of the Ottawa group and others who have examined ethical considerations in CRTs have broken new ground in that area. However, the analysis in the Ottawa statement has shortcomings when applied to IR. As argued here, their analysis of the equipoise requirement conflates IR with the more familiar type of RCT (or CRT) that tests new unproven interventions of various kinds. More generally, the boundary between IR and QI studies remains blurred, and the distinction can be made clearly only when research subjects or clusters are randomized. Similarly, the line between public health research and public health practice can be fuzzy when IR involves program evaluation. The tools, techniques and methodological rigor may be the same in both activities, but if the activity is construed as research, review by an REC is probably required. Nevertheless, as noted earlier, public health activities that are not research may pose ethical challenges that can benefit from ethics review. Researchers themselves need some guidance in these matters, as illustrated by the corrections the sponsoring organization's program officer had to make when reviewing the research protocols. Gray areas are likely to remain, but practical decisions must be made to ensure protection of the rights and welfare of human participants, be they patients, healthy volunteers or HCWs.

Acknowledgements

The author thanks the anonymous reviewers for their helpful suggestions.

Funding

This work was supported by the Brocher Foundation in Hermance, Switzerland, during the author's 1-month residency at the Foundation in February 2013. www.brocher.ch.

Notes

1. This example is drawn from an actual research protocol but altered slightly to protect confidentiality. I reviewed a number of cluster randomized protocols devoted to IR with permission granted by the sponsoring organization. Examination of these protocols contributed to my thinking about the topics addressed in this article.
2. Identity withheld to preserve confidentiality.
3. These distinctions illuminate the debate surrounding the proper way of denoting human beings in research either as 'subjects' or as 'participants.' While the HCWs are the subjects (their behavior is what is being studied), the patients are participants but not the subjects in this type of IR. 'Participant' is too vague a term to make the distinction in question here.

References

Centers for Disease Control and Prevention (2010). Distinguishing Public Health Research and Public Health Nonresearch. Available from: http://www.cdc.gov/od/science/integrity/docs/cdc-policy-distinguishing-public-health-research-nonresearch.pdf [accessed 6 February 2014].

Fairchild, A. L. (2003). Dealing with Humpty Dumpty: Research, Practice, and the Ethics of Public Health Surveillance. *The Journal of Law, Medicine & Ethics*, **31**, 615–623.

Lynn, J. (2004). When Does Quality Improvement Count as Research? Human Subject Protection and Theories of Knowledge. *Quality and Safety in Health Care*, **13**, 67–70.

McRae, A. D., Weijer, C., Binik, A., Grimshaw, J. M., Boruch, R., Brehaut, J. C., Donner, A., Eccles, M. P., Saginur, R., White, A. and Taljaard, M. (2011a). When Is Informed Consent Required in Cluster Randomized Trials in Health Research? *Trials*, **12,** 202. Available from: http://www.trialsjournal.com/content/12/1/202 [accessed 6 February 2014].

McRae, A. D., Weijer, C., Binik, A., White, A., Grimshaw, J. M., Baruch, R., Brehaut, J. C., Donner, A., Eccles, M., Saginur, R., Zwarenstein, M. and Taljaard, M. (2011b). Who Is the Research Subject in Cluster Randomized Trials in Health Research? *Trials*, **12,** 183. Available from: http://www.trialsjournal.com/content/12/1/183 [accessed 6 February 2014].

Peters, D. H., Tran, N. T. and Adam, T. (2013). Implementation Research in Health: A Practical Guide. Geneva: World Health Organization. Available from: http://www.who.int/alliance-hpsr/alliancehpsr_irpguide.pdf. [accessed 6 February 2014].

Remme, J. H. F., Taghreed, A., Becerra-Posada, F., D'Arcangues, C., Devlin, M., Gardner, C., Ghaffar, A., Hombach, J., Kengeya, J. F. K., Mbewu, A., Mbizvo, M. T., Mirza, Z., Pang, T., Ridley, R. G., Ziker, R. and Terry, R. F. (2010). Defining Research to Improve Health Systems. *PLoS Medicine*, 7:e1001000. Available from: http://www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.1001000 [accessed 6 February 2014].

Sim, J. and Dawson, A. (2012). Informed Consent and Cluster-Randomized Trials. *American Journal of Public Health*, **102**, 480–485.

Srinivasan, S. (2013). Ethics of 'standard care' in randomised controlled trials of screening for cervical cancer. *Indian Journal of Medical Ethics*, **10**, 147–149.

Weijer, C., Grimshaw, J. M., Eccles, M. P., McRae, A. D., White, A., Brehaut, J. C., Taljaard, M., The Ottawa Ethics of Cluster Randomized Trials Consensus Group and Althabe, F. (2012). The Ottawa Statement on the Ethical Design and Conduct of Cluster Randomized Trials. *PLoS medicine 9:* e1001346. Available at http://www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.1001346 [accessed 6 February 2014].

Weijer, C., Grimshaw, J. M., Taljaard, M., Binik, A., Boruch, R., Brehaut, J. C., Donner, A., Eccles, M. P., Gallo, A., McRae, A. D., Saginur, R. and Zwarenstein, M. (2011). Ethical Issues Posed by Cluster Randomized Trials in Health Research. *Trials*, 12: 100. Available from: http://www.trialsjournal.com/content/12/1/100. [accessed 6 February 2014].

World Health Organization. (2012) Cervical cancer screening in developing countries: report of a WHO consultation. Available from: http://www.who.int/cancer/publications/cervical_cancer_screening/en/. [accessed 6 February 2014].

World Health Organization. (2013) Implementation Research Platform. Available from: http://www.who.int/alliance-hpsr/projects/implementationresearch/en/ [accessed 6 February 2014].