Determing When Quality Improvement Initiatives Should Be Considered Research: Proposed Criteria and Potential Implications

David Casarett; Jason H. T. Karlawish; Jeremy Sugarman


http://jama.ama-assn.org/cgi/content/full/283/17/2275

Correction
Contact me if this article is corrected.

Citations
This article has been cited 66 times.
Contact me when this article is cited.

Topic collections
Medical Practice; Medical Practice, Other; Quality of Care; Quality of Care, Other; Statistics and Research Methods
Contact me when new articles are published in these topic areas.

Related Articles published in the same issue
May 3, 2000

Related Letters
Should Patients in Quality-Improvement Activities Have the Same Protections as Participants in Research Studies? Shan Cretin et al. JAMA. 2000;284(14):1786.
Determining When Quality Improvement Initiatives Should Be Considered Research Proposed Criteria and Potential Implications

David Casarett, MD, MA
Jason H. T. Karlawish, MD
Jeremy Sugarman, MD, MPH, MA

Recent developments in health care have been associated with a large number of initiatives aimed at improving the quality of health care. Many of these initiatives are described as quality improvement (QI). Quality improvement has become a major force in shaping health care and eliciting rapid change with the goal of improving the quality of health care provided by hospitals, clinics, and health care systems.¹

However, it is often unclear when QI initiatives are also research. This distinction is not simply a descriptive convenience. Whether a QI initiative is also classified as research may determine whether that effort is reviewed before it is initiated and what measures would be used to protect patients and others who participate. For instance, if an initiative is classified as research, an institutional review board (IRB) first must agree that any risks related to the initiative are reasonable with respect to the potential benefits and the importance of the knowledge to be gained.² Federal regulations also may require informed consent, perhaps with written documentation.²(116-117) However, if the same initiative is not research, such procedures are not usually required.

Incorrect classification may have significant implications. On the one hand, research initiatives that are misclassified solely as QI may be conducted without appropriate review, potentially resulting in harms to patients and others involved. For instance, a recent large-scale application of a new ultrasound screening technique for trisomy 21 was conducted as a QI initiative.³ However, in the ensuing debate, critics and medical journal editors debated whether this screening procedure was in fact research and, as such, whether it should have been reviewed by an IRB.⁴,⁵ On the other hand, if IRBs had to review all QI initiatives, institutions would be overly burdened with additional work in their efforts to improve care and protect patients. In addition, important QI initiatives might be unnecessarily hampered by procedures that are part of the current system for reviewing research.

In this article, we discuss some of the challenges of determining when QI initiatives should be reviewed and regulated as research. We examine the applicability of federal regulations governing this area and discuss other criteria that have been proposed. We then propose 2 stepwise criteria that might be used to identify QI initiatives that require review as research un-
Quality improvement initiatives may be difficult to distinguish from research because federal regulations broadly define research as “a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge.” \(^2\) Under this definition, research is tied to the investigators’ intent to produce generalizable knowledge.

Two difficulties occur when this definition is applied to QI initiatives. First, the intent of those engaged in QI may be difficult to determine. It may be easy to argue that the intent of a QI intervention is to produce information that would be applicable only locally. One means of assessing intent, often discussed as a heuristic technique, is that an initiative should be considered research if those conducting it hope to publish the results. \(^3\) However, although the intent to publish the results may be clear at the outset, such a decision is certainly open to change and is not an indirect measure of intent to create generalizable knowledge.

Second, a definition based on generalizability is difficult to apply because the results of virtually all QI interventions are generalizable, to lesser or greater degrees. For instance, although QI data would be of direct interest to those who participate, these data also may be generalizable to others in the same institution and perhaps to other institutions. Under the current federal definition of research, it is not clear whether the requirement to create generalizable information refers to knowledge that can be applied simply to other patients within an institution or beyond it.

Because a definition based solely on generalizability may be difficult to apply, 2 other criteria have been proposed. \(^4\) One criterion is the degree to which an initiative deviates from standard practice and another is whether an initiative requires identifiable recruitment practices. These criteria are logical and have considerable intuitive appeal. However, they do not help make a determination of whether an initiative meets the federal definition of research, and it is unclear how these factors should be weighted. For example, it is not clear how an initiative should be assessed if it produces generalizable knowledge but does not involve formal recruitment and does not deviate from standard practice.

**DETERMINING WHEN QI INITIATIVES SHOULD BE CONSIDERED RESEARCH**

Although such criteria may encourage detailed analysis of whether a QI initiative should be reviewed as research, they also may make the decisions of IRBs ambiguous and inconsistent. \(^5\) The challenge, therefore, is to develop criteria that are ethically justifiable and that can be applied consistently. We suggest that this can be accomplished by considering a QI initiative in light of 2 sequential criteria. First, a QI initiative should be reviewed and regulated as research if the majority of patients are not expected to benefit directly from the knowledge to be gained. If the majority of patients involved in a QI initiative are likely to benefit from the knowledge to be gained, a second criterion applies. That initiative should be reviewed and regulated as research if participants would be subjected to additional risks or burdens beyond usual clinical practice to make its results generalizable.

These criteria highlight the rights and interests of patients. Although QI initiatives also can involve clinicians, we have focused on protecting the rights and interests of patients because the protections available to patients and clinicians differ significantly. Patients are less able than are clinicians to protect...
their rights and interests. This is particularly true when patients are not even aware that a QI initiative exists or is ongoing. Conversely, clinicians have access to protections aside from IRB review, which include peer review organizations, negotiation with the QI project leader, and opting out of a health care system. Although clinicians are affected by QI, it is the patient who most needs the protection of an IRB review.

**Criterion 1: Direct Benefit to Patients Involved**

An initiative should be reviewed and regulated as research if the majority of patients are not expected to benefit directly from the knowledge to be gained. The question that operationalizes this criterion is, “Are the majority of patients expected to benefit from the knowledge to be gained?” Clinical practice is the benchmark by which to answer this question. Experienced clinicians at the site where the initiative is planned should agree that substantial evidence demonstrates that the majority of patients who participate would benefit from the knowledge to be gained. To make this assessment, clinicians should compare the initiative with standard clinical practice at that institution.

A QI initiative that does not meet this criterion must be reviewed and regulated as research. An initiative that does meet this criterion might not require review as research if a second criterion, discussed below, is met. In applying this first criterion, it seems appropriate to consider patients as individuals, rather than as members of a system of care, because this frame of analysis makes the application of current regulations for research manageable.

This first criterion requires that several kinds of initiatives be reviewed and regulated as research because the knowledge they generate would not benefit the patients involved (TABLE). For instance, an intervention to improve cost-effectiveness would produce knowledge that would not benefit the patients exposed to the intervention. Similarly, an initiative designed to develop prognostic criteria for patients in an intensive care unit is unlikely to produce knowledge relevant to those patients, unless they are readmitted to the same unit.

If the initiative involves a patient satisfaction survey administered to patients after hospital discharge, the knowledge generated would benefit only those patients who would be readmitted. Therefore, such an initiative should be reviewed as research unless those designing the initiative could demonstrate that a majority of patients would be readmitted and thus would benefit from the knowledge to be gained. Conversely, if the same satisfaction survey involved patients from a stable clinic population, in which a majority of patients could be expected to benefit from future application of the knowledge to be gained, this QI initiative might not require review as re-

<table>
<thead>
<tr>
<th>Example</th>
<th>Criterion 1: Direct Benefit to Patients Involved</th>
<th>Criterion 2: Additional Risks or Burdens</th>
<th>Reviewed as Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chart review of intensive care unit discharges</td>
<td>Knowledge will only benefit patients if it follows the patients to their next site of care</td>
<td>Not applicable</td>
<td>Activity is research based on criterion 1 alone</td>
</tr>
<tr>
<td>Cost-effectiveness analysis of a physician reminder system to decrease lengths of stay</td>
<td>Patient may benefit from shorter lengths of stay but not from knowledge about the cost-effectiveness of this intervention</td>
<td>Not applicable</td>
<td>Activity is research based on criterion 1 alone</td>
</tr>
<tr>
<td>Postdischarge satisfaction survey</td>
<td>Patients unlikely to benefit from the knowledge to be gained unless knowledge is applicable to future site of care</td>
<td>Not applicable</td>
<td>Activity is research based on criterion 1 alone</td>
</tr>
<tr>
<td>Clinic satisfaction survey with stable patient population</td>
<td>Majority of patients could be expected to benefit from the information to be gained</td>
<td>No</td>
<td>No additional risks or burdens are imposed to make results generalizable beyond this population</td>
</tr>
<tr>
<td>Clinic patients randomized to receive appointment reminders, and no-show rate measured</td>
<td>Stable clinic patients would benefit from information to be gained</td>
<td>No</td>
<td>Randomization is not standard care but poses no risks; no other additional risks or burdens are imposed</td>
</tr>
<tr>
<td>Comparison of 2 approved urinary catheters in a chronic care unit, with weekly catheter changes and urine cultures</td>
<td>Population is stable; patients involved will benefit from knowledge to be gained</td>
<td>Yes</td>
<td>Generalizability requires additional data and risk or burdens (catheter changes)</td>
</tr>
</tbody>
</table>
search. In all of these examples, predictions about follow-up and the likelihood of benefit are required. These predictions may not always be easy, but they are no more difficult than are similar predictions about benefit and follow-up that are made routinely by clinicians.

This first criterion has 3 caveats. First, there is room for considerable debate about the proportion of patients that are expected to benefit to make a determination about whether a QI initiative should be reviewed and regulated as research. For instance, a result of a screening study should be applicable to the care of most of the patients involved. Similarly, in a patient satisfaction survey, the majority of patients should stand to benefit from changes made on the basis of their responses. This threshold has its ethical justification in what most patients would likely find appropriate, namely being included in initiatives that lead to improved health care. While assuming that the majority threshold is reasonable, we recognize it might be modified in some instances based on empirical data about the preferences of clinicians, investigators, and patients.

Second, this criterion assesses whether the patients being evaluated would benefit directly from the knowledge to be gained, not whether they would benefit from participating. Although the latter is important in considering the ethics of clinical care and proposed research, the benefit of potential knowledge to be gained should guide a determination about whether a QI initiative should be reviewed and regulated as research. For instance, patients in a randomized trial of appointment reminders may benefit if they receive a reminder and if these reminders are effective. However, to fulfill the first criterion, the patients who participate in the initiative must likely benefit from knowledge about whether the reminder system is effective.

Third, the first criterion does not determine whether informed consent is needed, because informed consent is necessary both for research and clinical care. Consequently, the decision about whether a QI initiative is reviewed as research should not depend on the administrative burden of obtaining informed consent. For example, it is likely that meeting informed consent requirements for research may make a QI initiative more difficult when institutions or hospital wards are randomized. This may motivate those who design these initiatives to consider them solely as QI to avoid these requirements, but such an approach motivated by convenience is clearly inappropriate. If an initiative is really research, it should be reviewed and regulated as such, keeping in mind that federal research regulations permit waivers of informed consent for certain types of research. However, the decision to waive informed consent for research should be made by an IRB, only after it is determined that an initiative constitutes research.

In summary, the first criterion stipulates that if a QI initiative produces knowledge that would directly benefit the majority of patients involved, it may not need to be reviewed and regulated as research. Conversely, if the QI initiative does not, it should be reviewed as research. Of course, many initiatives produce knowledge that is applicable both to the patients under study and other future patients. Even though clinical trials are designed to produce generalizable knowledge, the participants in those trials also may benefit from knowledge about the medications under study. This suggests the need for another criterion.

**Criterion 2: Imposition of Additional Risks or Burdens**

A QI initiative should be reviewed and regulated as research if it imposes risks or burdens beyond the standard of practice to make the results generalizable. Under this criterion, the central question is, “Will generalizability require additional risks or burdens?” To answer this question, the judgment of experienced clinicians defines the standard of practice. These clinicians should be familiar with the disease or condition and with current local practice at the site where the initiative is being planned.

Risks or burdens that are imposed to make results generalizable might be created by additional tests or surveys or by additional data collection strategies that could result in a loss of privacy or confidentiality. For example, if a clinic sent appointment reminders to half its patients and then measured the rate of missed appointments in each group, no additional risks or burdens need to be imposed to make the results generalizable. However, if additional information were gathered from patients’ medical records or from the patients’ clinicians to characterize better the patient population under study, the initiative poses additional risks and burdens and therefore should be reviewed and regulated as research.

Similarly, a randomized comparison of 2 clinically approved urethral catheters for patients residing in an extended care facility and requiring catheterization might not need to be treated as research if no additional risks or burdens are imposed to make the results generalizable. However, if the same study required more catheter changes than standard clinical practice, it should be treated as research. Many might object to this conclusion and would argue that the random assignment of patients, hospital wards, or institutions is “clearly” research. We agree that this opinion is intuitively plausible and that similar to the heuristic involving the intent to publish the results of a QI initiative, it may be a useful heuristic to decide that an intervention is research. However, it is unclear how this position might be ethically justified.

Although randomization in and of itself may not require that an initiative be considered research, informed consent still should be obtained when the initiative involves interventions, such as use of a urethral catheter, that require clinical informed consent. Specifically, in addition to information about the catheters being used, patients should understand that they are being randomly assigned to 1 of 2 interventions. Nonetheless, in most cases,
we believe that when patients are randomly assigned, additional data usually would be gathered to make the results more generalizable. For instance, a study that randomly assigns patients to receive 1 of 2 urinary catheters most likely would gather additional data to make the results generalizable. Review of medical records or additional urinary cultures may pose additional risks or burdens to the patient, which would require that the study be reviewed and regulated as research. Therefore, although it is possible to envision a randomized study that is not research, this should be rare.

To answer the question “Will generalizability require additional risks or burdens?” the only risks that are relevant are those that exceed standard clinical practice and that are imposed to make the results generalizable. The second criterion adapts what Freedman and colleagues have described as “demarcated research risk.” This is the principle that research risks should be assessed in relation to the usual care that patients would receive if they were not part of a study. Therefore, what is central to determining whether an initiative should be considered research is whether additional risks and burdens are imposed to make the results generalizable.

Thus, this second criterion requires an assessment of intent for 1 or more parts of the initiative, such as a chart review, an increased frequency of catheter changes, or an additional survey instrument. This criterion does not apply to the intent of the initiative as a whole. The intent of the initiative as a whole may be difficult to define, but the intent of a single aspect of the initiative can be examined and more easily defined.

Implications of the Proposed Criteria

These 2 proposed criteria should provide much needed guidance to IRBs, clinicians, and administrators about when QI initiatives should be reviewed and regulated as research. These criteria also should assist medical journal editors, who have agreed that IRB approval is a necessary condition for publication, to determine whether to publish the results of initiatives that are described as QI. Taken together, these 2 criteria offer guidance in determining when initiatives should be reviewed as research. This does not mean that applying these 2 criteria always will be straightforward. These criteria require judgments about whether the majority of patients would benefit from the knowledge to be gained or whether generalizability requires additional risks or burdens to participants. However, these criteria provide a foundation for deliberation about whether a QI initiative should be reviewed as research.

These criteria offer appropriate incentives for those who design QI initiatives. Research requires IRB review and typically involves elaborate informed consent procedures, which can be cumbersome, if not prohibitive. Intentionally, these requirements may lead those planning or conducting such initiatives to avoid the designation of “research” by producing knowledge that is of limited applicability. Instead, these 2 criteria offer incentives that are more consistent with the goals of science and clinical care. For instance, the first criterion offers an incentive to ensure that knowledge gained from a study is applied directly to the care of the patients exposed to the intervention. The second criterion offers an additional incentive to minimize risks and burdens to those that are necessary to provide knowledge that would benefit those individuals who participate.

Despite their potential advantages, these 2 criteria may pose considerable challenges to institutions because the resulting number of QI initiatives that could be reviewed as research could be extensive. As an example, cost-effectiveness analyses produce knowledge that is unlikely to benefit the patients involved. Therefore, these studies would be evaluated as research unless they meet the federal requirements for an exemption from review. An exemption might be appropriate if participants are not identifiable, the information is already publicly available, or the study is designed to evaluate a public service or benefit program under federal jurisdiction. Nevertheless, if these criteria were adopted widely, the result might well increase the workload of IRBs.

One response to this observation would be to argue that the criteria themselves should be rejected on the grounds that they would create unprecedented burdens on health care institutions and IRBs. However, we believe that it makes little sense to reject these criteria, if they are otherwise sound, simply because they would create additional burdens for institutions. Therefore, if these criteria are to be applied without undue burden on investigators and IRBs, efficient mechanisms of review would need to be established.

At least 3 possibilities exist. First, the standard procedures of IRB review, with accepted criteria for exemptions (eg, waiver of informed consent), might be required of all such initiatives. Second, a requirement for IRB review might be limited to IRB notification about initiatives that constitute exempt research under federal regulations, obviating a need for complete review and approval. Third, institutions that are engaged in frequent QI initiatives may find it helpful to establish a committee that is responsible for reviewing QI projects. All of these options have advantages and disadvantages and may be appropriate in different settings, depending on the unique needs of various institutions and their patients.

CONCLUSIONS

These 2 criteria and efficient mechanisms to apply them should help provide guidelines for determining whether a QI effort should be reviewed and regulated as research. But equally important is the need for close scrutiny of the processes of developing, implementing, and evaluating QI efforts. Indeed, one of the reasons that this problem is so difficult is that while guidelines exist for responsible research and clinical care, no such guidelines exist for QI.
Medicine and society have developed clear processes of research and clinical care. Research studies have a principal investigator (PI) who is responsible for the proper conduct of the research. A corpus of codes, regulations, and guidelines defines the proper conduct of that PI. A culture of training and peer review as well as IRB review exists to ensure the PI conducts research properly. Similarly, the standards that guide clinical practice are equally well defined and include medical training, certification, continuing education, meeting legal standards for care, and obtaining informed consent.

The problem is that QI falls somewhere between these 2 practices, both conceptually and administratively. Conceptually, QI is strictly neither clinical care nor research, and perhaps resembles a public health initiative more than either of these. In QI initiatives, the responsibility of those planning and conducting the initiative is neither to individual patients nor to science, but rather to populations of patients, often grouped within an institution.

The administrative features of QI also are quite different. Most illustrative of this difference, perhaps, is that project directors for QI initiatives may not be as easily identifiable, or as visible to patients, as are clinicians or the PI of a research study. Clinicians or PIs are readily identifiable and are held responsible for clinical care and research, respectively. Their identities are known to IRBs, funding agencies, and usually to the patients who enroll in a study. However, responsibility for QI initiatives may be more difficult to assign. Ultimately, what might be needed is a system of ethical oversight that can guide institutions that engage in QI initiatives. This ethical system should begin at a local level by developing ways to apply the 2 criteria we propose herein. But any effective system of oversight also would need to define ways in which QI initiatives should be conducted responsibly. At a minimum, this would include maintaining confidentiality, minimizing burdens to participants, and procedures for managing data. Future discussion should extend to the national level to ensure that all patients can rely on equal protection. Until these rules for conduct can be agreed on and disseminated, a gap will continue to exist between the scrutiny that is applied to QI vs research and clinical care. When QI has protections akin to those of clinical care and research and when there is general confidence about the responsible conduct of QI, the decision about when QI initiatives should be reviewed and regulated as research will become less important.

REFERENCES
17. Salgo v Leland Stanford University, 317 P2d 150 (Calif 1957).