

Position Paper ■

The Use and Interpretation of Quasi-Experimental Studies in Medical Informatics

ANTHONY D. HARRIS, MD, MPH, JESSINA C. MCGREGOR, PhD, ELI N. PERENCEVICH, MD, MS, JON P. FURUNO, PhD, JINGKUN ZHU, MS, DAN E. PETERSON, MD, MPH, JOSEPH FINKELSTEIN, MD

Abstract Quasi-experimental study designs, often described as nonrandomized, pre-post intervention studies, are common in the medical informatics literature. Yet little has been written about the benefits and limitations of the quasi-experimental approach as applied to informatics studies. This paper outlines a relative hierarchy and nomenclature of quasi-experimental study designs that is applicable to medical informatics intervention studies. In addition, the authors performed a systematic review of two medical informatics journals, the *Journal of the American Medical Informatics Association* (JAMIA) and the *International Journal of Medical Informatics* (IJMI), to determine the number of quasi-experimental studies published and how the studies are classified on the above-mentioned relative hierarchy. They hope that future medical informatics studies will implement higher level quasi-experimental study designs that yield more convincing evidence for causal links between medical informatics interventions and outcomes.

■ *J Am Med Inform Assoc.* 2006;13:16–23. DOI 10.1197/jamia.M1749.

Background

Quasi-experimental studies encompass a broad range of non-randomized intervention studies. These designs are frequently used when it is not logistically feasible or ethical to conduct a randomized controlled trial. Examples of quasi-experimental studies follow. As one example of a quasi-experimental study, a hospital introduces a new order-entry system and wishes to study the impact of this intervention on the number of medication-related adverse events before and after the intervention. As another example, an informatics technology group is introducing a pharmacy order-entry system aimed at decreasing pharmacy costs. The intervention is implemented and pharmacy costs before and after the intervention are measured.

In medical informatics, the quasi-experimental, sometimes called the pre-post intervention, design often is used to evaluate the benefits of specific interventions. The increasing capacity of health care institutions to collect routine clinical data has led to the growing use of quasi-experimental study

designs in the field of medical informatics as well as in other medical disciplines. However, little is written about these study designs in the medical literature or in traditional epidemiology textbooks.^{1–3} In contrast, the social sciences literature is replete with examples of ways to implement and improve quasi-experimental studies.^{4–6}

In this paper, we review the different pretest-posttest quasi-experimental study designs, their nomenclature, and the relative hierarchy of these designs with respect to their ability to establish causal associations between an intervention and an outcome. The example of a pharmacy order-entry system aimed at decreasing pharmacy costs will be used throughout this article to illustrate the different quasi-experimental designs. We discuss limitations of quasi-experimental designs and offer methods to improve them. We also perform a systematic review of four years of publications from two informatics journals to determine the number of quasi-experimental studies, classify these studies into their application domains, determine whether the potential limitations of quasi-experimental studies were acknowledged by the authors, and place these studies into the above-mentioned relative hierarchy.

Methods

The authors reviewed articles and book chapters on the design of quasi-experimental studies.^{4–10} Most of the reviewed articles referenced two textbooks that were then reviewed in depth.^{4,6}

Key advantages and disadvantages of quasi-experimental studies, as they pertain to the study of medical informatics, were identified. The potential methodological flaws of quasi-experimental medical informatics studies, which have the potential to introduce bias, were also identified. In addition, a summary table outlining a relative hierarchy and nomenclature of quasi-experimental study designs is described. In general, the higher the design is in the hierarchy, the greater the internal validity that the study traditionally

Affiliations of the authors: Department of Epidemiology and Preventive Medicine, University of Maryland, Baltimore, MD (ADH, JCM, ENP, JPF, JZ, JF); VA Maryland Health Care System, Baltimore, MD (ADH, ENP); Cereplex Inc., Germantown, MD (DEP).

Dr. Harris was supported by NIH grants K23 AI01752-01A1 and R01 AI60859-01A1. Dr. Perencevich was supported by a VA Health Services Research and Development Service (HSR&D) Research Career Development Award (RCD-02026-1). Dr. Finkelstein was supported by NIH grant RO1 HL71690.

Correspondence and reprints: Anthony D. Harris, MD, MPH, Division of Healthcare Outcomes Research, Department of Epidemiology and Preventive Medicine, University of Maryland School of Medicine, 100 N. Greene Street, Lower Level, Baltimore, MD; e-mail: <aharris@epi.umaryland.edu>.

Received for review: 11/19/04; accepted for publication: 08/12/05.

possesses because the evidence of the potential causation between the intervention and the outcome is strengthened.⁴

We then performed a systematic review of four years of publications from two informatics journals. First, we determined the number of quasi-experimental studies. We then classified these studies on the above-mentioned hierarchy. We also classified the quasi-experimental studies according to their application domain. The categories of application domains employed were based on categorization used by Yearbooks of Medical Informatics 1992–2005 and were similar to the categories of application domains employed by Annual Symposiums of the American Medical Informatics Association.¹¹ The categories were (1) health and clinical management; (2) patient records; (3) health information systems; (4) medical signal processing and biomedical imaging; (5) decision support, knowledge representation, and management; (6) education and consumer informatics; and (7) bioinformatics. Because the quasi-experimental study design has recognized limitations, we sought to determine whether authors acknowledged the potential limitations of this design. Examples of acknowledgment included mention of lack of randomization, the potential for regression to the mean, the presence of temporal confounders and the mention of another design that would have more internal validity.

All original scientific manuscripts published between January 2000 and December 2003 in the *Journal of the American Medical Informatics Association* (JAMIA) and the *International Journal of Medical Informatics* (IJMI) were reviewed. One author (ADH) reviewed all the papers to identify the number of quasi-experimental studies. Other authors (ADH, JCM, JF) then independently reviewed all the studies identified as quasi-experimental. The three authors then convened as a group to resolve any disagreements in study classification, application domain, and acknowledgment of limitations.

Results and Discussion

What Is a Quasi-experiment?

Quasi-experiments are studies that aim to evaluate interventions but that do not use randomization. Similar to randomized trials, quasi-experiments aim to demonstrate causality between an intervention and an outcome. Quasi-experimental studies can use both preintervention and postintervention measurements as well as nonrandomly selected control groups.

Using this basic definition, it is evident that many published studies in medical informatics utilize the quasi-experimental design. Although the randomized controlled trial is generally considered to have the highest level of credibility with regard to assessing causality, in medical informatics, researchers often choose not to randomize the intervention for one or more reasons: (1) ethical considerations, (2) difficulty of randomizing subjects, (3) difficulty to randomize by locations (e.g., by wards), (4) small available sample size. Each of these reasons is discussed below.

Ethical considerations typically will not allow random withholding of an intervention with known efficacy. Thus, if the efficacy of an intervention has not been established, a randomized controlled trial is the design of choice to determine efficacy. But if the intervention under study incorporates an accepted, well-established therapeutic intervention, or if the

intervention has either questionable efficacy or safety based on previously conducted studies, then the ethical issues of randomizing patients are sometimes raised. In the area of medical informatics, it is often believed prior to an implementation that an informatics intervention will likely be beneficial and thus medical informaticians and hospital administrators are often reluctant to randomize medical informatics interventions. In addition, there is often pressure to implement the intervention quickly because of its believed efficacy, thus not allowing researchers sufficient time to plan a randomized trial.

For medical informatics interventions, it is often difficult to randomize the intervention to individual patients or to individual informatics users. So while this randomization is technically possible, it is underused and thus compromises the eventual strength of concluding that an informatics intervention resulted in an outcome. For example, randomly allowing only half of medical residents to use pharmacy order-entry software at a tertiary care hospital is a scenario that hospital administrators and informatics users may not agree to for numerous reasons.

Similarly, informatics interventions often cannot be randomized to individual locations. Using the pharmacy order-entry system example, it may be difficult to randomize use of the system to only certain locations in a hospital or portions of certain locations. For example, if the pharmacy order-entry system involves an educational component, then people may apply the knowledge learned to nonintervention wards, thereby potentially masking the true effect of the intervention. When a design using randomized locations is employed successfully, the locations may be different in other respects (confounding variables), and this further complicates the analysis and interpretation.

In situations where it is known that only a small sample size will be available to test the efficacy of an intervention, randomization may not be a viable option. Randomization is beneficial because on average it tends to evenly distribute both known and unknown confounding variables between the intervention and control group. However, when the sample size is small, randomization may not adequately accomplish this balance. Thus, alternative design and analytical methods are often used in place of randomization when only small sample sizes are available.

What Are the Threats to Establishing Causality When Using Quasi-experimental Designs in Medical Informatics?

The lack of random assignment is the major weakness of the quasi-experimental study design. Associations identified in quasi-experiments meet one important requirement of causality since the intervention precedes the measurement of the outcome. Another requirement is that the outcome can be demonstrated to vary statistically with the intervention. Unfortunately, statistical association does not imply causality, especially if the study is poorly designed. Thus, in many quasi-experiments, one is most often left with the question: "Are there alternative explanations for the apparent causal association?" If these alternative explanations are credible, then the evidence of causation is less convincing. These rival hypotheses, or alternative explanations, arise from principles of epidemiologic study design.

Shadish et al.⁴ outline nine threats to internal validity that are outlined in Table 1. Internal validity is defined as the degree to which observed changes in outcomes can be correctly inferred to be caused by an exposure or an intervention. In quasi-experimental studies of medical informatics, we believe that the methodological principles that most often result in alternative explanations for the apparent causal effect include (a) difficulty in measuring or controlling for important *confounding variables*, particularly unmeasured confounding variables, which can be viewed as a subset of the *selection* threat in Table 1; (b) results being explained by the statistical principle of *regression to the mean*. Each of these latter two principles is discussed in turn.

An inability to sufficiently control for important confounding variables arises from the lack of randomization. A variable is a confounding variable if it is associated with the exposure of interest and is also associated with the outcome of interest; the confounding variable leads to a situation where a causal association between a given exposure and an outcome is observed as a result of the influence of the confounding variable. For example, in a study aiming to demonstrate that the introduction of a pharmacy order-entry system led to lower pharmacy costs, there are a number of important potential confounding variables (e.g., severity of illness of the patients, knowledge and experience of the software users, other changes in hospital policy) that may have differed in the pre-intervention and postintervention time periods (Fig. 1). In a multivariable regression, the first confounding variable could be addressed with severity of illness measures, but the second confounding variable would be difficult if not nearly impossible to measure and control. In addition, potential confounding variables that are unmeasured or immeasurable cannot be controlled for in nonrandomized quasi-experimental study designs and can only be properly controlled by the randomization process in randomized controlled trials.

Another important threat to establishing causality is regression to the mean.¹²⁻¹⁴ This widespread statistical phenomenon can result in wrongly concluding that an effect is due to the intervention when in reality it is due to chance. The phenomenon was first described in 1886 by Francis Galton

Table 1 ■ Threats to Internal Validity

1. Ambiguous temporal precedence: Lack of clarity about whether intervention occurred before outcome
2. Selection: Systematic differences over conditions in respondent characteristics that could also cause the observed effect
3. History: Events occurring concurrently with intervention could cause the observed effect
4. Maturation: Naturally occurring changes over time could be confused with a treatment effect
5. Regression: When units are selected for their extreme scores, they will often have less extreme subsequent scores, an occurrence that can be confused with an intervention effect
6. Attrition: Loss of respondents can produce artifactual effects if that loss is correlated with intervention
7. Testing: Exposure to a test can affect scores on subsequent exposures to that test
8. Instrumentation: The nature of a measurement may change over time or conditions
9. Interactive effects: The impact of an intervention may depend on the level of another intervention

Adapted from Shadish et al.⁴

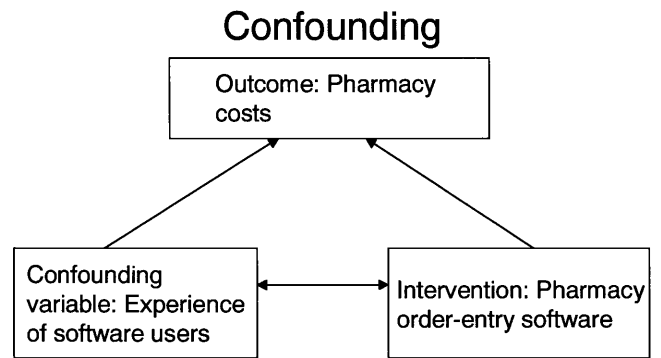


Figure 1. Example of confounding. To get the true effect of the intervention of interest, we need to control for the confounding variable.

who measured the adult height of children and their parents. He noted that when the average height of the parents was greater than the mean of the population, the children tended to be shorter than their parents, and conversely, when the average height of the parents was shorter than the population mean, the children tended to be taller than their parents.

In medical informatics, what often triggers the development and implementation of an intervention is a rise in the rate above the mean or norm. For example, increasing pharmacy costs and adverse events may prompt hospital informatics personnel to design and implement pharmacy order-entry systems. If this rise in costs or adverse events is really just an extreme observation that is still within the normal range of the hospital's pharmaceutical costs (i.e., the mean pharmaceutical cost for the hospital has not shifted), then the statistical principle of regression to the mean predicts that these elevated rates will tend to decline even without intervention. However, often informatics personnel and hospital administrators cannot wait passively for this decline to occur. Therefore, hospital personnel often implement one or more interventions, and if a decline in the rate occurs, they may mistakenly conclude that the decline is causally related to the intervention. In fact, an alternative explanation for the finding could be regression to the mean.

What Are the Different Quasi-experimental Study Designs?

In the social sciences literature, quasi-experimental studies are divided into four study design groups^{4,6}:

- A. Quasi-experimental designs without control groups
- B. Quasi-experimental designs that use control groups but no pretest
- C. Quasi-experimental designs that use control groups and pretests
- D. Interrupted time-series designs

There is a relative hierarchy within these categories of study designs, with category D studies being sounder than categories C, B, or A in terms of establishing causality. Thus, if feasible from a design and implementation point of view, investigators should aim to design studies that fall in to the higher rated categories. Shadish et al.⁴ discuss 17 possible designs, with seven designs falling into category A, three designs in category B, and six designs in category C, and one major design in category D. In our review, we determined that most medical informatics quasi-experiments could be

characterized by 11 of 17 designs, with six study designs in category A, one in category B, three designs in category C, and one design in category D because the other study designs were not used or feasible in the medical informatics literature. Thus, for simplicity, we have summarized the 11 study designs most relevant to medical informatics research in Table 2.

The nomenclature and relative hierarchy were used in the systematic review of four years of JAMIA and the IJMI. Similar to the relative hierarchy that exists in the evidence-based literature that assigns a hierarchy to randomized controlled trials, cohort studies, case-control studies, and case series, the hierarchy in Table 2 is not absolute in that in some cases, it may be infeasible to perform a higher level

Table 2 ■ Relative Hierarchy of Quasi-experimental Designs

Quasi-experimental Study Designs	Design Notation
A. Quasi-experimental designs without control groups	
1. The one-group posttest-only design	X O1
2. The one-group pretest-posttest design	O1 X O2
3. The one-group pretest-posttest design using a double pretest	O1 O2 X O3
4. The one-group pretest-posttest design using a nonequivalent dependent variable	(O1a, O1b) X (O2a, O2b)
5. The removed-treatment design	O1 X O2 O3 removeX O4
6. The repeated-treatment design	O1 X O2 removeX O3 X O4
B. Quasi-experimental designs that use a control group but no pretest	
1. Posttest-only design with nonequivalent groups	Intervention group: X O1 Control group: O2
C. Quasi-experimental designs that use control groups and pretests	
1. Untreated control group with dependent pretest and posttest samples	Intervention group: O1a X O2a Control group: O1b O2b
2. Untreated control group design with dependent pretest and posttest samples using a double pretest	Intervention group: O1a O2a X O3a Control group: O1b O2b O3b
3. Untreated control group design with dependent pretest and posttest samples using switching replications	Intervention group: O1a X O2a O3a Control group: O1b O2b X O3b
D. Interrupted time-series design*	
1. Multiple pretest and posttest observations spaced at equal intervals of time	O1 O2 O3 O4 O5 X O6 O7 O8 O9 O10

O = Observational Measurement; X = Intervention Under Study. Time moves from left to right.

*In general, studies in category D are of higher study design quality than studies in category C, which are higher than those in category B, which are higher than those in category A. Also, as one moves down within each category, the studies become of higher quality, e.g., study 5 in category A is of higher study design quality than study 4, etc.

study. For example, there may be instances where an A6 design established stronger causality than a B1 design.¹⁵⁻¹⁷

Quasi-experimental Designs without Control Groups

The One-Group Posttest-Only Design (X O1)

Here, X is the intervention and O is the outcome variable (this notation is continued throughout the article). In this study design, an intervention (X) is implemented and a posttest observation (O1) is taken. For example, X could be the introduction of a pharmacy order-entry intervention and O1 could be the pharmacy costs following the intervention. This design is the weakest of the quasi-experimental designs that are discussed in this article. Without any pretest observations or a control group, there are multiple threats to internal validity. Unfortunately, this study design is often used in medical informatics when new software is introduced since it may be difficult to have pretest measurements due to time, technical, or cost constraints.

The One-Group Pretest-Posttest Design (O1 X O2)

This is a commonly used study design. A single pretest measurement is taken (O1), an intervention (X) is implemented, and a posttest measurement is taken (O2). In this instance, period O1 frequently serves as the "control" period. For example, O1 could be pharmacy costs prior to the intervention, X could be the introduction of a pharmacy order-entry system, and O2 could be the pharmacy costs following the intervention. Including a pretest provides some information about what the pharmacy costs would have been had the intervention not occurred.

The One-Group Pretest-Posttest Design Using a Double Pretest (O1 O2 X O3)

The advantage of this study design over A2 is that adding a second pretest prior to the intervention helps provide evidence that can be used to refute the phenomenon of regression to the mean and confounding as alternative explanations for any observed association between the intervention and the posttest outcome. For example, in a study where a pharmacy order-entry system led to lower pharmacy costs ($O3 < O2$ and $O1$), if one had two preintervention measurements of pharmacy costs (O1 and O2) and they were both elevated, this would suggest that there was a decreased likelihood that O3 is lower due to confounding and regression to the mean. Similarly, extending this study design by increasing the number of measurements postintervention could also help to provide evidence against confounding and regression to the mean as alternate explanations for observed associations.

The One-Group Pretest-Posttest Design Using a Non-equivalent Dependent Variable ([O1a, O1b] X [O2a, O2b])

This design involves the inclusion of a nonequivalent dependent variable (*b*) in addition to the primary dependent variable (*a*). Variables *a* and *b* should assess similar constructs; that is, the two measures should be affected by similar factors and confounding variables except for the effect of the intervention. Variable *a* is expected to change because of the intervention X, whereas variable *b* is not. Taking our example, variable *a* could be pharmacy costs and variable *b* could be the length of stay of patients. If our informatics intervention is aimed at decreasing pharmacy costs, we would expect to observe a

decrease in pharmacy costs but not in the average length of stay of patients. However, a number of important confounding variables, such as severity of illness and knowledge of software users, might affect both outcome measures. Thus, if the average length of stay did not change following the intervention but pharmacy costs did, then the data are more convincing than if just pharmacy costs were measured.

The Removed-Treatment Design (O1 X O2 O3 Remove X O4)

This design adds a third posttest measurement (O3) to the one-group pretest-posttest design and then removes the intervention before a final measure (O4) is made. The advantage of this design is that it allows one to test hypotheses about the outcome in the presence of the intervention and in the absence of the intervention. Thus, if one predicts a decrease in the outcome between O1 and O2 (after implementation of the intervention), then one would predict an increase in the outcome between O3 and O4 (after removal of the intervention). One caveat is that if the intervention is thought to have persistent effects, then O4 needs to be measured after these effects are likely to have disappeared. For example, a study would be more convincing if it demonstrated that pharmacy costs decreased after pharmacy order-entry system introduction (O2 and O3 less than O1) and that when the order-entry system was removed or disabled, the costs increased (O4 greater than O2 and O3 and closer to O1). In addition, there are often ethical issues in this design in terms of removing an intervention that may be providing benefit.

*The Repeated-Treatment Design
(O1 X O2 Remove X O3 X O4)*

The advantage of this design is that it demonstrates reproducibility of the association between the intervention and the outcome. For example, the association is more likely to be causal if one demonstrates that a pharmacy order-entry system results in decreased pharmacy costs when it is first introduced and again when it is reintroduced following an interruption of the intervention. As for design A5, the assumption must be made that the effect of the intervention is transient, which is most often applicable to medical informatics interventions. Because in this design, subjects may serve as their own controls, this may yield greater statistical efficiency with fewer numbers of subjects.

Quasi-experimental Designs That Use a Control Group but No Pretest

Posttest-Only Design with Nonequivalent Groups:

$$\frac{\text{Interventiongroup: } X \ O1}{\text{Controlgroup: } O2}$$

An intervention X is implemented for one group and compared to a second group. The use of a comparison group helps prevent certain threats to validity including the ability to statistically adjust for confounding variables. Because in this study design, the two groups may not be equivalent (assignment to the groups is not by randomization), confounding may exist. For example, suppose that a pharmacy order-entry intervention was instituted in the medical intensive care unit (MICU) and not the surgical intensive care unit (SICU). O1 would be pharmacy costs in the MICU after the intervention and O2 would be pharmacy costs in the SICU after the intervention. The absence of a pretest makes it difficult to know whether a change has occurred in the

MICU. Also, the absence of pretest measurements comparing the SICU to the MICU makes it difficult to know whether differences in O1 and O2 are due to the intervention or due to other differences in the two units (confounding variables).

Quasi-experimental Designs That Use Control Groups and Pretests

The reader should note that with all the studies in this category, the intervention is not randomized. The control groups chosen are comparison groups. Obtaining pretest measurements on both the intervention and control groups allows one to assess the initial comparability of the groups. The assumption is that if the intervention and the control groups are similar at the pretest, the smaller the likelihood there is of important confounding variables differing between the two groups.

Untreated Control Group with Dependent Pretest and Posttest Samples:

$$\frac{\text{Interventiongroup: } O1a \ X \ O2a}{\text{Controlgroup: } O1b \ O2b}$$

The use of both a pretest and a comparison group makes it easier to avoid certain threats to validity. However, because the two groups are nonequivalent (assignment to the groups is not by randomization), selection bias may exist. Selection bias exists when selection results in differences in unit characteristics between conditions that may be related to outcome differences. For example, suppose that a pharmacy order-entry intervention was instituted in the MICU and not the SICU. If preintervention pharmacy costs in the MICU (O1a) and SICU (O1b) are similar, it suggests that it is less likely that there are differences in the important confounding variables between the two units. If MICU postintervention costs (O2a) are less than preintervention MICU costs (O1a), but SICU costs (O1b) and (O2b) are similar, this suggests that the observed outcome may be causally related to the intervention.

Untreated Control Group Design with Dependent Pretest and Posttest Samples Using a Double Pretest:

$$\frac{\text{Interventiongroup: } O1a \ O2a \ X \ O3a}{\text{Controlgroup: } O1b \ O2b \ O3b}$$

In this design, the pretests are administered at two different times. The main advantage of this design is that it controls for potentially different time-varying confounding effects in the intervention group and the comparison group. In our example, measuring points O1 and O2 would allow for the assessment of time-dependent changes in pharmacy costs, e.g., due to differences in experience of residents, preintervention between the intervention and control group, and whether these changes were similar or different.

Untreated Control Group Design with Dependent Pretest and Posttest Samples Using Switching Replications:

$$\frac{\text{Interventiongroup: } O1a \ X \ O2a \ O3a}{\text{Controlgroup: } O1b \ O2b \ X \ O3b}$$

With this study design, the researcher administers an intervention at a later time to a group that initially served as a non-intervention control. The advantage of this design over design C2 is that it demonstrates reproducibility in two different settings. This study design is not limited to two groups; in fact, the study results have greater validity if the intervention effect is replicated in different groups at multiple times. In the

example of a pharmacy order-entry system, one could implement or intervene in the MICU and then at a later time, intervene in the SICU. This latter design is often very applicable to medical informatics where new technology and new software is often introduced or made available gradually.

Interrupted Time-Series Designs

O1 O2 O3 O4 O5 X O6 O7 O8 O9 O10

An interrupted time-series design is one in which a string of consecutive observations equally spaced in time is interrupted by the imposition of a treatment or intervention. The advantage of this design is that with multiple measurements both pre- and postintervention, it is easier to address and control for confounding and regression to the mean. In addition, statistically, there is a more robust analytic capability, and there is the ability to detect changes in the slope or intercept as a result of the intervention in addition to a change in the mean values.¹⁸ A change in intercept could represent an immediate effect while a change in slope could represent a gradual effect of the intervention on the outcome. In the example

of a pharmacy order-entry system, O1 through O5 could represent monthly pharmacy costs preintervention and O6 through O10 monthly pharmacy costs post the introduction of the pharmacy order-entry system. Interrupted time-series designs also can be further strengthened by incorporating many of the design features previously mentioned in other categories (such as removal of the treatment, inclusion of a nondependent outcome variable, or the addition of a control group).

Systematic Review Results

The results of the systematic review are in Table 3. In the four-year period of JAMIA publications that the authors reviewed, 25 quasi-experimental studies among 22 articles were published. Of these 25, 15 studies were of category A, five studies were of category B, two studies were of category C, and no studies were of category D. Although there were no studies of category D (interrupted time-series analyses), three of the studies classified as category A had data collected that could have been analyzed as an interrupted time-series analysis. Nine of the 25 studies (36%) mentioned at least one of the

Table 3 ■ Systematic Review of Four Years of Quasi-designs in JAMIA

Study	Journal	Informatics Topic Category	Quasi-experimental Design	Limitation of Quasi-design Mentioned in Article
Staggers and Kobus ²⁰	JAMIA	1	Counterbalanced study design	Yes
Schriger et al. ²¹	JAMIA	1	A5	Yes
Patel et al. ²²	JAMIA	2	A5 (study 1, phase 1)	No
Patel et al. ²²	JAMIA	2	A2 (study 1, phase 2)	No
Borowitz ²³	JAMIA	1	A2	No
Patterson and Harasym ²⁴	JAMIA	6	C1	Yes
Rocha et al. ²⁵	JAMIA	5	A2	Yes
Lovis et al. ²⁶	JAMIA	1	Counterbalanced study design	No
Hersh et al. ²⁷	JAMIA	6	B1	No
Makoul et al. ²⁸	JAMIA	2	B1	Yes
Ruland ²⁹	JAMIA	3	B1	No
DeLusignan et al. ³⁰	JAMIA	1	A1	No
Mekhjian et al. ³¹	JAMIA	1	A2 (study design 1)	Yes
Mekhjian et al. ³¹	JAMIA	1	B1 (study design 2)	Yes
Ammenwerth et al. ³²	JAMIA	1	A2	No
Oniki et al. ³³	JAMIA	5	C1	Yes
Liederman and Morefield ³⁴	JAMIA	1	A1 (study 1)	No
Liederman and Morefield ³⁴	JAMIA	1	A2* (study 2)	No
Rotich et al. ³⁵	JAMIA	2	A2*	No
Payne et al. ³⁶	JAMIA	1	A1	No
Hoch et al. ³⁷	JAMIA	3	A2*	No
Laerum et al. ³⁸	JAMIA	1	B1	Yes
Devine et al. ³⁹	JAMIA	1	Counterbalanced study design	
Dunbar et al. ⁴⁰	JAMIA	6	A1	
Lenert et al. ⁴¹	JAMIA	6	A2	
Koide et al. ⁴²	IJMI	5	D4	No
Gonzalez-Hendrich et al. ⁴³	IJMI	2	A1	No
Anantharaman and Swee Han ⁴⁴	IJMI	3	B1	No
Chae et al. ⁴⁵	IJMI	6	A2	No
Lin et al. ⁴⁶	IJMI	3	A1	No
Mikulich et al. ⁴⁷	IJMI	1	A2	Yes
Hwang et al. ⁴⁸	IJMI	1	A2	Yes
Park et al. ⁴⁹	IJMI	1	C2	No
Park et al. ⁴⁹	IJMI	1	D4	No

JAMIA = Journal of the American Medical Informatics Association; IJMI = International Journal of Medical Informatics.

*Could have been analyzed as an interrupted time-series design.

potential limitations of the quasi-experimental study design. In the four-year period of IJMI publications reviewed by the authors, nine quasi-experimental studies among eight manuscripts were published. Of these nine, five studies were of category A, one of category B, one of category C, and two of category D. Two of the nine studies (22%) mentioned at least one of the potential limitations of the quasi-experimental study design.

In addition, three studies from JAMIA were based on a counterbalanced design. A counterbalanced design is a higher order study design than other studies in category A. The counterbalanced design is sometimes referred to as a Latin-square arrangement. In this design, all subjects receive all the different interventions but the order of intervention assignment is not random.¹⁹ This design can only be used when the intervention is compared against some existing standard, for example, if a new PDA-based order entry system is to be compared to a computer terminal-based order entry system. In this design, all subjects receive the new PDA-based order entry system and the old computer terminal-based order entry system. The counterbalanced design is a within-participants design, where the order of the intervention is varied (e.g., one group is given software A followed by software B and another group is given software B followed by software A). The counterbalanced design is typically used when the available sample size is small, thus preventing the use of randomization. This design also allows investigators to study the potential effect of ordering of the informatics intervention.

Conclusion

Although quasi-experimental study designs are ubiquitous in the medical informatics literature, as evidenced by 34 studies in the past four years of the two informatics journals, little has been written about the benefits and limitations of the quasi-experimental approach. As we have outlined in this paper, a relative hierarchy and nomenclature of quasi-experimental study designs exist, with some designs being more likely than others to permit causal interpretations of observed associations. Strengths and limitations of a particular study design should be discussed when presenting data collected in the setting of a quasi-experimental study. Future medical informatics investigators should choose the strongest design that is feasible given the particular circumstances.

References ■

- Rothman KJ, Greenland S. *Modern epidemiology*. Philadelphia: Lippincott-Raven Publishers, 1998.
- Hennekens CH, Buring JE. *Epidemiology in medicine*. Boston: Little, Brown, 1987.
- Szklo M, Nieto FJ. *Epidemiology: beyond the basics*. Gaithersburg, MD: Aspen Publishers, 2000.
- Shadish WR, Cook TD, Campbell DT. *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin, 2002.
- Trochim WMK. *The research methods knowledge base*. Cincinnati: Atomic Dog Publishing, 2001.
- Cook TD, Campbell DT. *Quasi-experimentation: design and analysis issues for field settings*. Chicago: Rand McNally Publishing Company, 1979.
- MacLehose RR, Reeves BC, Harvey IM, Sheldon TA, Russell IT, Black AM. A systematic review of comparisons of effect sizes derived from randomised and non-randomised studies. *Health Technol Assess*. 2000;4:1-154.
- Shadish WR, Heinsman DT. Experiments versus quasi-experiments: do they yield the same answer? *NIDA Res Monogr*. 1997;170:147-64.
- Grimshaw J, Campbell M, Eccles M, Steen N. Experimental and quasi-experimental designs for evaluating guideline implementation strategies. *Fam Pract*. 2000;17(Suppl 1):S11-6.
- Zwerling C, Daltroy LH, Fine LJ, Johnston JJ, Melius J, Silverstein BA. Design and conduct of occupational injury intervention studies: a review of evaluation strategies. *Am J Ind Med*. 1997;32:164-79.
- Haux RKC, editor. *Yearbook of medical informatics 2005*. Stuttgart: Schattauer Verlagsgesellschaft, 2005, p. 563.
- Morton V, Torgerson DJ. Effect of regression to the mean on decision making in health care. *BMJ*. 2003;326:1083-4.
- Bland JM, Altman DG. Regression towards the mean. *BMJ*. 1994;308:1499.
- Bland JM, Altman DG. Some examples of regression towards the mean. *BMJ*. 1994;309:780.
- Guyatt GH, Haynes RB, Jaeschke RZ, Cook DJ, Green L, Naylor CD, et al. Users' guides to the medical literature: XXV. Evidence-based medicine: principles for applying the users' guides to patient care. Evidence-Based Medicine Working Group. *JAMA*. 2000;284:1290-6.
- Harris RP, Helfand M, Woolf SH, Lohr KN, Mulrow CD, Teutsch SM, et al. Current methods of the US Preventive Services Task Force: a review of the process. *Am J Prev Med*. 2001;20:21-35.
- Harbour R, Miller J. A new system for grading recommendations in evidence based guidelines. *BMJ*. 2001;323:334-6.
- Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. Segmented regression analysis of interrupted time series studies in medication use research. *J Clin Pharm Ther*. 2002;27:299-309.
- Campbell DT. Counterbalanced design. In: Company RMCP, editor. *Experimental and Quasiexperimental Designs for Research*. Chicago: Rand-McNally College Publishing Company, 1963, p. 50-5.
- Staggers N, Kobus D. Comparing response time, errors, and satisfaction between text-based and graphical user interfaces during nursing order tasks. *J Am Med Inform Assoc*. 2000;7:164-76.
- Schriger DL, Baraff LJ, Buller K, Shendrikar MA, Nagda S, Lin EJ, et al. Implementation of clinical guidelines via a computer charting system: effect on the care of febrile children less than three years of age. *J Am Med Inform Assoc*. 2000;7:186-95.
- Patel VL, Kushniruk AW, Yang S, Yale JF. Impact of a computer-based patient record system on data collection, knowledge organization, and reasoning. *J Am Med Inform Assoc*. 2000;7:569-85.
- Borowitz SM. Computer-based speech recognition as an alternative to medical transcription. *J Am Med Inform Assoc*. 2001;8:101-2.
- Patterson R, Harasym P. Educational instruction on a hospital information system for medical students during their surgical rotations. *J Am Med Inform Assoc*. 2001;8:111-6.
- Rocha BH, Christenson JC, Evans RS, Gardner RM. Clinicians' response to computerized detection of infections. *J Am Med Inform Assoc*. 2001;8:117-25.
- Lovis C, Chapko MK, Martin DP, Payne TH, Baud RH, Hoey PJ, et al. Evaluation of a command-line parser-based order entry pathway for the Department of Veterans Affairs electronic patient record. *J Am Med Inform Assoc*. 2001;8:486-98.
- Hersh WR, Junium K, Mailhot M, Tidmarsh P. Implementation and evaluation of a medical informatics distance education program. *J Am Med Inform Assoc*. 2001;8:570-84.
- Makoul G, Curry RH, Tang PC. The use of electronic medical records: communication patterns in outpatient encounters. *J Am Med Inform Assoc*. 2001;8:610-5.
- Ruland CM. Handheld technology to improve patient care: evaluating a support system for preference-based care planning at the bedside. *J Am Med Inform Assoc*. 2002;9:192-201.

30. De Lusignan S, Stephens PN, Adal N, Majeed A. Does feedback improve the quality of computerized medical records in primary care? *J Am Med Inform Assoc.* 2002;9:395-401.
31. Mekhjian HS, Kumar RR, Kuehn L, Bentley TD, Teater P, Thomas A, et al. Immediate benefits realized following implementation of physician order entry at an academic medical center. *J Am Med Inform Assoc.* 2002;9:529-39.
32. Ammenwerth E, Mansmann U, Iller C, Eichstadter R. Factors affecting and affected by user acceptance of computer-based nursing documentation: results of a two-year study. *J Am Med Inform Assoc.* 2003;10:69-84.
33. Oniki TA, Clemmer TP, Pryor TA. The effect of computer-generated reminders on charting deficiencies in the ICU. *J Am Med Inform Assoc.* 2003;10:177-87.
34. Liederman EM, Morefield CS. Web messaging: a new tool for patient-physician communication. *J Am Med Inform Assoc.* 2003;10:260-70.
35. Rotich JK, Hannan TJ, Smith FE, Bii J, Odero WW, Vu N, Mamlin BW, et al. Installing and implementing a computer-based patient record system in sub-Saharan Africa: the Mosoriot Medical Record System. *J Am Med Inform Assoc.* 2003;10:295-303.
36. Payne TH, Hoey PJ, Nichol P, Lovis C. Preparation and use of preconstructed orders, order sets, and order menus in a computerized provider order entry system. *J Am Med Inform Assoc.* 2003;10:322-9.
37. Hoch I, Heymann AD, Kurman I, Valinsky LJ, Chodick G, Shalev V. Countrywide computer alerts to community physicians improve potassium testing in patients receiving diuretics. *J Am Med Inform Assoc.* 2003;10:541-6.
38. Laerum H, Karlsen TH, Faxvaag A. Effects of scanning and eliminating paper-based medical records on hospital physicians' clinical work practice. *J Am Med Inform Assoc.* 2003;10:588-95.
39. Devine EG, Gaehde SA, Curtis AC. Comparative evaluation of three continuous speech recognition software packages in the generation of medical reports. *J Am Med Inform Assoc.* 2000;7:462-8.
40. Dunbar PJ, Madigan D, Grohskopf LA, Revere D, Woodward J, Minstrell J, et al. A two-way messaging system to enhance anti-retroviral adherence. *J Am Med Inform Assoc.* 2003;10:11-5.
41. Lenert L, Munoz RF, Stoddard J, Delucchi K, Bansod A, Skoczen S, et al. Design and pilot evaluation of an Internet smoking cessation program. *J Am Med Inform Assoc.* 2003;10:16-20.
42. Koide D, Ohe K, Ross-Degnan D, Kaihara S. Computerized reminders to monitor liver function to improve the use of etretinate. *Int J Med Inf.* 2000;57:11-9.
43. Gonzalez-Heydrich J, DeMaso DR, Irwin C, Steingard RJ, Kohane IS, Beardslee WR. Implementation of an electronic medical record system in a pediatric psychopharmacology program. *Int J Med Inf.* 2000;57:109-16.
44. Anantharaman V, Swee Han L. Hospital and emergency ambulance link: using IT to enhance emergency pre-hospital care. *Int J Med Inf.* 2001;61:147-61.
45. Chae YM, Heon Lee J, Hee Ho S, Ja Kim H, Hong Jun K, Uk Won J. Patient satisfaction with telemedicine in home health services for the elderly. *Int J Med Inf.* 2001;61:167-73.
46. Lin CC, Chen HS, Chen CY, Hou SM. Implementation and evaluation of a multifunctional telemedicine system in NTUH. *Int J Med Inf.* 2001;61:175-87.
47. Mikulich VJ, Liu YC, Steinfeldt J, Schriger DL. Implementation of clinical guidelines through an electronic medical record: physician usage, satisfaction and assessment. *Int J Med Inf.* 2001;63:169-78.
48. Hwang JI, Park HA, Bakken S. Impact of a physician's order entry (POE) system on physicians' ordering patterns and patient length of stay. *Int J Med Inf.* 2002;65:213-23.
49. Park WS, Kim JS, Chae YM, Yu SH, Kim CY, Kim SA, et al. Does the physician order-entry system increase the revenue of a general hospital? *Int J Med Inf.* 2003;71:25-32.