21st Century Civil Justice System

A Roadmap for Reform

Measuring Innovation

National Center for State Courts

Institute for the Advancement of the American Legal System

University of Denver

21st Century Civil Justice System:

A Roadmap for Reform

Measuring Innovation

Institute for the Advancement of the American Legal System

at the

University of Denver

National Center for State Courts

UNIVERSITY OF DENVER | INSTITUTE *for the* ADVANCEMENT *of the* AMERICAN LEGAL SYSTEM

2044 E. Evans Avenue, HRTM Bldg., #307 Denver, CO 80208-2101
Phone: 303.871.6600
www.du.edu/legalinstitute

The Institute for the Advancement of the American Legal System (IAALS) is a national, non-partisan organization dedicated to improving the process and culture of the civil justice system. Executive Director and former Colorado Supreme Court Justice Rebecca Love Kourlis leads a staff distinguished not only by their expertise but also by their diversity of ideas, backgrounds and beliefs.

IAALS provides principled leadership, conducts comprehensive and objective research, and develops innovative and practical solutions—all focused on serving the individuals and organizations who rely on the system to clarify rights and resolve disputes.

IAALS is part of the University of Denver, and it benefits from gifts donated to the University for its use. None of those gifts have conditions or requirements, other than accounting and fiduciary responsibility. All IAALS research and products are supported by pooled grants from individuals, businesses and private foundations.

Rebecca Love Kourlis, Executive Director

Stephen Daniels, Director of Research

Pamela A. Gagel, Assistant Director

John Greacen, Consultant

300 Newport Avenue Williamsburg, VA 23185
Phone: 800.616.6164
www.ncsc.org

The National Center for State Courts (NCSC), headquartered in Williamsburg, VA, is a nonprofit organization dedicated to improving the administration of justice by providing leadership and service to the state courts and courts around the world. Founded in 1971 by the Conference of Chief Justices and Chief Justice of the United States Warren E. Burger, NCSC provides education, training, technology, management, and research services to the nation's state courts. The NCSC also has offices in Denver, CO, Arlington, VA, and Washington, DC.

Mary C. McQueen, President

Richard Schauffler, Director of Research Services

Paula Hannaford-Agor, Principal Court Research Consultant

Nicole Waters, Senior Court Research Associate

Contents

Executive Summary 2-3

Evaluating Civil Justice Reforms 4-9

Reasonable People May Predict

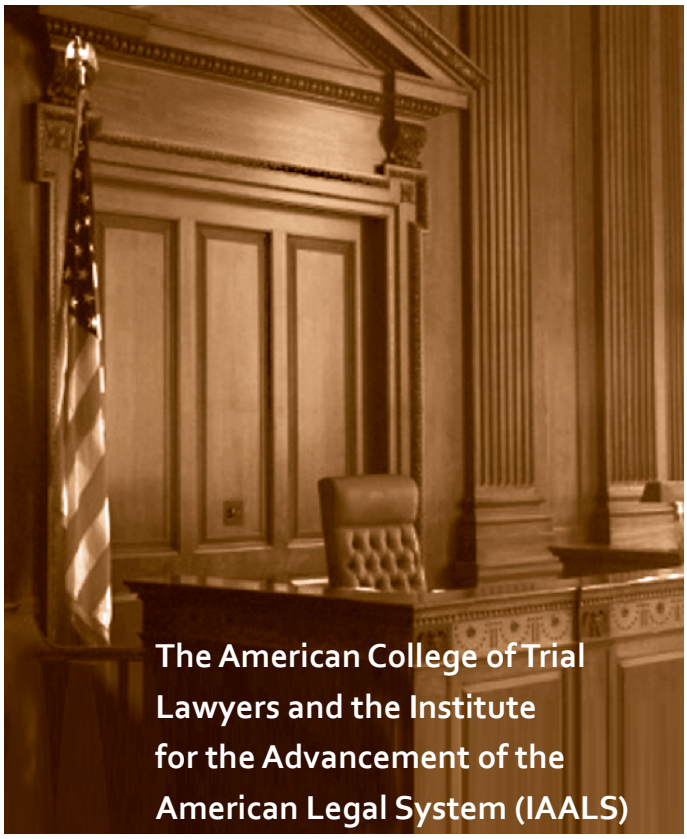
- Different Outcomes 10-13

Data To Be Collected 14-15

- Court Case Management Data 14
- Supplemental Data 14
- Survey, Focus Group, and Interview Data 15
- Monitoring Cases and Outcomes 15

Conclusion 16

Executive Summary



The American College of Trial Lawyers and the Institute for the Advancement of the American Legal System (IAALS) have promulgated proposals for changes to the rules of civil procedure as they currently exist in most United States jurisdictions and IAALS has published a guide for improving civil caseflow management practices. These suggested rules and caseflow management practices are available for any jurisdiction that wishes to implement them in whole or in part.

Anticipating that these developments will spur a significant number of federal and state courts to implement changes in their procedures for civil cases, IAALS and the National Center for State Courts (NCSC) have collaborated on an approach to evaluation of civil justice reforms that draws on the experience of both organizations in evaluating court programs and court performance. The methodology described here builds upon the NCSC's experience over the past thirty years with the Trial Court Performance Standards and more recently *CourTools*, a set of ten performance measures designed for performance measurement and performance management in the state trial courts. Further, this approach draws on the quality cycle process for continuous court improvement described in the NCSC's High Performance Courts Framework.

For any jurisdiction choosing to implement a civil rules or civil caseflow management reform project, the IAALS and the NCSC strongly urge an early and consistent focus on evaluation and performance measurement. This focus is necessary in order to test the theories on which the reforms are based and to identify ways in which to improve their implementation. Improving civil justice is not an event, but a process. Evaluation allows courts to

see whether the reforms are producing the intended outcomes, and, if not, to identify additional steps required for refining the reforms and measuring the outcomes again. This cycle of continuous improvement is the most effective management strategy for improving access to fair and timely resolution of civil cases.

This document provides an overview of a recommended approach to evaluation and describes how civil justice rules and caseflow management reforms should be evaluated. Many civil justice reform efforts have failed due to the fact that the court has not consistently implemented the reforms adopted in its jurisdiction, allowing parties to stipulate to exceptions to the rules to the point that the exceptions have swallowed the rules. Even where courts have been more successful in this aspect of implementation, too often they have failed to define, collect and analyze the data needed to provide an objective picture of the actual effects of the reforms implemented.

A key feature of the IAALS/NCSC approach is that the outcome measures associated with the reforms must be discussed and agreed at the *beginning* of the process. Courts choosing to implement civil justice reform projects must understand the necessity and value of consistent implementation of new procedures and of collecting empirical data regarding the actual results of any changes adopted. The evaluation should be planned at the outset so that appropriate analyses can be designed to allow the data to illuminate key outcomes of the reform process. These results can serve as the foundation for recommendations to the court concerning the future course of the reform project in that jurisdiction, while also allowing other courts around the nation to benefit from the experiences of the pilot court.

For courts throughout the country to obtain maximum benefit from empirical analyses of the results of civil rule and caseflow management changes, it is crucially important that all courts implementing this evaluation methodology collect the same or very similar data and that the analysis of that data employ the same definitions, counting rules, measurement techniques, and analytical strategies. We urge courts to consult with NCSC staff to ensure that their evaluation efforts comply with this principle.

This methodology describes in general terms the processes that can be used to examine empirically the impact of changes to civil rules and civil caseflow management principles. The expected outcomes associated with civil reforms are specified, along with how each should be measured. With this overview, courts can prepare to undertake civil justice reforms with the knowledge of what is required to evaluate and refine them.

Evaluating Civil Justice Reforms

Determining the effects of a procedural change requires *comparison* of court performance data with and without the new procedure. Merely reporting the performance of the court with the new procedure in place is not sufficient to determine whether the procedure has improved or diminished the court's effectiveness.

A variety of comparative methods exist and this document describes three of the most common. We urge courts to adopt one of the following three approaches and we explain the benefits and limitations of each: 1) comparing the court's performance before and after the change is made (known as a "pre-post" research design); 2) comparing the performance of similar courts, some of which implement the change while others maintain the previous process (known as a "comparison group" research design); 3) and comparing the performance of judges or judicial departments within the same court, some of whom adopt the change while others continue with business as usual (known as an "pre-post with comparison group" research design). The latter is the preferred approach, although we recognize it will not always be possible to implement this design.

The table at right summarizes these three approaches to evaluation design.

Evaluation Design	Data/Source		Data/Source	Comment
Pre-post	Baseline data from pilot courts/judges ***before*** the rules reforms are implemented	**are compared with**	Outcome data from pilot courts/judges ***after*** the rules reforms are implemented	Cannot distinguish effects of reforms from effects of external factors
Comparison group	Outcome data from comparison group courts/judges ***after*** the rules reforms are implemented	**are compared with**	Outcome data from pilot courts/judges ***after*** the rules reforms are implemented	Cannot distinguish effects of reforms from effects produced by differences among courts
Pre-post with comparison group	Baseline data from pilot and comparison courts/judges ***before*** the rules reforms are implemented	**are compared with**	Outcome data from pilot and comparison courts/judges ***after*** the rules reforms are implemented	Accounts for both external factors and differences among courts by comparing change over time between pilot and comparison courts

Pre-post Design In a pre-post study, an organization gathers *baseline* data on its operations before the new procedures are implemented and compares those data with data on its operations after the new procedures have been in place for a sufficient time for their effects to become apparent. For instance, a court could measure the average time from filing to disposition of civil cases resolved in the 12 months prior to the commencement of a pilot program and again during the 24 months after the new procedures were put in place. This plan would be based on an assumption that the full results of the new procedures would take at least two years before they would become apparent, given the fact that many civil cases take at least that period of time to reach resolution.

A common mistake in a pre-post study is to proceed with the new reforms without actually gathering the baseline data—merely assuming that the necessary data is available from the court's automated case management information system or from other sources. In fact, the work involved in correctly organizing the currently available data will identify any required data elements that are missing from the current information systems and inform the court of any supplemental data gathering that may be required before the new procedures are implemented.

The main benefit of a pre-post research design is that it allows comparison before and after an event (civil rules reform) has taken place. Measuring outcomes before the reform avoids the error of simply ascribing any changes in outcomes after the reform to the reform itself. While it sounds logical to think that any result that takes place after the reform itself is the product of the reform, this is not so. It is possible that other factors—such as characteristics of the cases, parties, judges, attorneys, or other factors in the legal or broader political, social, or economic environment—will also change during the period the reforms are implemented and that these factors will also affect the outcomes under evaluation.

Comparison Group Design In a comparison group research design, the changes are implemented

in only some of the courts in a state or in a few federal courts (the *experimental* group). Other non-implementing jurisdictions (also known as the *control* group) are identified that are as similar as possible to the implementing jurisdictions—for instance, similar in the population of the districts, the size of the courts' civil caseloads, the nature of the civil cases filed, the number of judges and staff assigned to civil cases, and the existing civil caseflow management approaches. Data are collected in both the implementing (*experimental*) and the non-implementing (*control*) jurisdictions and compared.

The results of a program evaluation that uses a comparison group design are often viewed as more persuasive, providing credible evidence for the results of the program. This is due to the fact that utilizing a comparison group design reflects an awareness of and attempt to control for characteristics that are important to the evaluation. Comparison group design is a good, practical approach to taking into account key factors that will influence the outcomes under study. At the same time, it should be understood, for example, that no two jurisdictions or judges or cases are truly identical, thus finding truly comparable controls is an inherently imperfect science.

Pre-post with Comparison Group Design In a pre-post with comparison group research design, multiple comparisons are made possible. Both the *experimental* and *control* groups can be compared to their own performance *before* and *after* the reforms were implemented. In addition, the two groups can be compared to each other. This combination of comparisons, when possible, provides the most systematic way to evaluate the effects of the reforms.

The principal challenge of a pre-post with comparison group study is to ensure the assignment of cases into the two groups in such a way that does not introduce bias. This can be achieved by seeking to ensure that the courts, judges, and cases participating in the study are as comparable as possible. The ideal solution to this challenge is to utilize some form of random

assignment. Random assignment means that cases are assigned to either the control group or the experimental group without regard to observed or unobserved case characteristics, such that any case has an equal probability of being assigned to either group. In this way, any observed differences in outcomes can be correctly attributed to the intervention (i.e., the new rules or procedures).

However, it is often not practical to utilize this approach. Thus, the best real world alternative is often to carefully select the courts and judges participating in the study, to seek to ensure that a representative group of cases are heard in all participating courts, and be alert to the fact that some observed or unobserved differences between the outcomes for cases proceeding through the new process may be what in fact produce the difference in outcomes, rather than the changed rules or procedures. Courts should guard against the temptation to pick the "good cases" (according to whatever criteria are thought to make them better candidates for the experimental group) or the "good judges" for the experiment because they want the reform program to be seen as a success. This effectively defeats any ability to objectively evaluate the reforms. Similarly, allowing parties or attorneys wide latitude to "opt in" or "opt out" of a new process has the same effect.

Thus, the success of this evaluation methodology rests on the ability of the court to ensure that there are no systematic differences in the cases handled under the old and new processes—that is, between *control* and *experimental* cases. This means that the court will need to ensure that cases are randomly assigned to the judges using the old and new processes and that the parties and their lawyers are not able to "steer" or "direct" cases from one category (e.g., the old process) to the other (e.g., the new process). The court should anticipate that individual lawyers will perceive an advantage for a particular case (or for all their cases) to be in one or the other category and will attempt to manipulate the system to achieve that result if humanly possible. (This manipulation process is often referred to as "judge shopping" even though it may in fact more

properly be named "process shopping" in civil justice reform projects.)

With a comparison group design, only some of the civil judges will be able to implement the new process. The other judges will have to continue to follow the old process in order to produce data for the control cases. It will be essential that the control group judges not allow themselves—intentionally or unintentionally—to begin to follow any part of the new process, or to begin to adopt the "spirit" of the new process. Comparison of the results of the control group to its own baseline data in the pre-post with comparison group design will allow for an additional point of comparison that may suggest whether the control group judges have actually changed their practices. Conversely, experimental group judges should not allow parties to stipulate to deviations from the reformed rules and procedures for the same reason.

Whatever the evaluation design chosen, the key is to conduct the evaluation in a consistent manner. Program evaluations, even using the most optimal design, generally cannot be said to "prove" that a program works or not. To do that would require a true experiment, which is not practical in the real world. The evaluation can, however, provide credible evidence that can be readily understood by a wide audience that a civil rules reform or caseflow management program is making a difference.

Reasonable People May Predict Different Outcomes

In discussions of civil rules reform, a variety of anticipated improvements are cited. For some of these improvements, consensus may exist on the expected outcome of the particular reform (e.g., time to disposition will be faster). For others, judges, attorneys, and court staff may make reasonable arguments for differing, even opposing outcomes (e.g., "fewer trials will take place and more cases will be disposed earlier" vs. "more trials will take place because reformed rules will make it possible for more cases to get to trial"). It would be very difficult for anyone to predict in advance what all the possible combinations of effects might be; some reforms may push in one direction, while others push in the opposite direction. The only way to know is to measure.

Measurement of expected outcomes from any changes to rules and procedures (e.g., "more trials") sounds simple enough, but it requires three key elements that all courts participating in the evaluation should follow. First, definitions of key terms (e.g., what is a case? what is a trial?) and key data elements (e.g., filing date, disposition date, appearance, case type) are required that are clear, uniform, and unambiguous. Courts participating in an evaluation should discuss and agree in advance how they will define and count key outcome measures. For example, some courts count a disposition as a "jury trial" disposition at any point after a jury trial is requested and a trial date is assigned; some count it as such if a jury has been sworn; and still others count it that way only after a jury has been sworn and the first evidence introduced. Second, for all courts, the counting rules must be defined and followed consistently. For example, some courts overcount cases by counting a case with multiple parties or causes of action as many "cases" rather than seeing this as a single case. Some courts stop the time-to-disposition clock when a civil case is halted due to the initiation of a bankruptcy proceeding in federal court by one of the involved parties, while others do not. For an evaluation to be successful, the definitions and counting rules need to be used consistently in measuring outcomes. The best way to ensure this is to embed those definitions and rules in the court's case management system and train court staff to use the data entry codes consistently. If automation is not an option, training staff and monitoring the manual collection of data is critical to collecting data that can accurately measure results.

Finally, the effect of rules reform on the outcomes described below may vary by case type, so a proper evaluation will include differentiation of cases by case type (e.g., medical malpractice, product liability, automobile torts, subrogation) in order to fully appreciate

the impact of new civil procedures. Conforming the case types of the court to the case types defined in the NCSC's *State Court Guide to Statistical Reporting* will ensure maximum comparability across jurisdictions.

The examples below illustrate how outcomes can be defined and matched to performance measures. A key feature of the IAALS/NCSC approach is that these outcome measures must be agreed upon as part of the discussion about rules reform at the *beginning* of the process, with agreement as to what will be measured and how. Confirmation of the availability of data prior to commencement of implementation assures a consistent and well-balanced evaluation will be possible. For each adopted rule change or case management change, discussion should be held to make explicit the expected outcomes and their measures. Illustrative examples of expected outcomes and their respective measures include:

The number of civil trials is expected to increase/decrease IAALS and NCSC note that there is a reasonable difference of opinion among civil practitioners about the probable impact of civil rules reforms on the trial rate. Many believe that rule revisions that apply proportionality to civil procedures based on the value and complexity of the case will result in more cases progressing to trial because the path to trial has been made shorter and less expensive, while others believe that simplified and proportional procedures will result in fewer trials, because more cases will settle in advance of a trial.

Performance measures:

- Number of civil cases disposed of by bench trial
- Number of civil cases disposed of by jury trial
- The trial rate
- Number of civil cases disposed of by non-trial dispositions

The number of expert witnesses per case will decrease Rules reform is aimed at limiting the number of experts retained by the parties during discovery as well as the number of experts actually called at the time of trial.

Performance measures:

- Number of expert witnesses used by each side during discovery
- Number of expert witnesses used by each side during trial

The number of motions per case will decrease

The number of discovery motions will be of particular interest, but other motions are important as well because of their cost implications.

Performance measures:

- Number of motions per case
- Number of discovery motions per case

Reasonable People May Predict Different Outcomes

The number of discovery events per case will decrease Because discovery documents are not typically filed with the court, it will be necessary to require the parties to report to the court when they serve a set of interrogatories, a request for documents, or a deposition notice.

Performance measures:

- Average number of sets of interrogatories per case
- Average number of requests for documents per case
- Average number of depositions per case (expert, non-expert)
- Average length of depositions

The length of trials is expected to increase/decrease Some civil justice reforms are designed in part to shorten the length of trials, as procedures are made more proportional to the complexity of the case and are designed to encourage maximum disclosure of known issues early in the process. Whether this is true is subject to empirical validation.

Performance measures:

- The length of bench trials
- The length of jury trials

The number of events per case is expected to increase/decrease Consistent with their differences of opinion on the effect of civil justice reforms on trial rates, civil justice experts do not agree on whether the new procedures will entail more or fewer court appearances for the parties.

Performance measure:

- Number of events per case, by event type

The time to disposition is expected to decrease Fewer case events, less time between case events, and simplified discovery are expected to decrease the overall time to disposition. The research design should take into account the frequency of and elapsed time between intermediate case events (e.g., time from filing of complaint to filing of answer, to first hearing, to filing of a motion for summary judgment, to decision on a motion for summary judgment, to a pretrial conference, and to a trial) that may affect time to disposition. Because courts will have different events, and will call them by different names, it will be necessary to develop a way of assigning court events to standardized event categories.

Performance measures:

- Time from filing a case to its disposition, in days
- Percentage of cases resolved within defined time standards
- Age of active pending caseload
- Time between key events

- Time from filing a motion to resolution of the issues raised in that motion, in days

Levels of litigant and lawyer satisfaction

For litigants, shorter time to disposition, lower costs, and increased access to the court will increase satisfaction. A number of factors may increase lawyer satisfaction. Civil rules reform will mean an ability to meet the needs of their clients in a reasonable manner and to ensure each case receives the attention it deserves and finds a forum in the court. Rules reform may also constrain overly aggressive legal tactics and promote greater collegiality and a higher level of intellectual satisfaction with the quality of lawyering.

Performance measures:

- Procedural fairness and satisfaction surveys for litigants
- Satisfaction survey for lawyers

Costs of litigation are expected to decrease

The NCSC is developing an approach to cost-modeling that is designed to estimate litigation costs based on the number of case events, and the frequency and duration of those events, controlling for case type and local jurisdictional factors. This cost model will allow meaningful comparisons between estimated costs pre- and post- civil rules reform, as well as across jurisdictions. In addition, the costs can be decomposed into the cost of phases of a case or events within a case (e.g., discovery).

The lesser number of case events, fewer experts and interrogatories, and shorter time frames contemplated by civil rules reform should result in lower cost per case. However, it is also the case that as rules change, behavior changes to compensate. For example, restrictions on the number of experts may result in experts charging higher fees, or fact pleading may result in increased up-front costs, which may or may not be offset by decreased costs in later stages of a case. Some of these effects may not be seen or felt for some time, while others may be relatively immediate. For these reasons, measuring cost per case is an important component of civil rules reform evaluation.

Performance measures:

- The estimated cost per case to parties and to the court using the NCSC event-based cost model

These examples of important outcomes and appropriate performance measures are merely illustrative, and should be augmented by issues of particular interest to the implementing court and its bar. The aim is to design an evaluation that is sufficiently complex to meaningfully capture the reality it is trying to describe yet also simple enough to be practical and possible to implement.

Data To Be Collected

The data to be collected will be specific to the particular rule and procedural changes implemented in the pilot court. As indicated previously, each performance measure requires specific, well-defined data elements and consistent analytical methods to produce meaningful results.

Court Case Management Data Most of these data will be generated by the court's case management system; some courts or states export and house these data in a data warehouse that can serve as the data source. It is essential that for each performance measure undertaken, the court invests the time to examine the supporting data. This is essential both for establishing the baseline data and for ensuring that the pre-post and comparison group data is collected accurately.

Where examination of existing data reveals deficiencies in the information collected, the court must ensure that these deficiencies are corrected in advance of the implementation of any reform. Court staff may need additional training in the proper use of specific event codes and proper data entry for specific events and rulings over the life of a case. A classic problem in this regard is the overuse of generic event codes, the all-inclusive "Other" field in the case management system, or inappropriate overuse of text fields for recording specific events.

Realistically, a court may not be able to reengineer its case management system for the purpose of this evaluation. However, it is also true that information systems can produce far more information than they are typically used for, if a court can specify what it requires in its management reports. An investment in exploring this option on the front end of the project can save significant amounts of time and effort on the back end. In addition, automating the reporting of this information is the best way to ensure consistency in correctly classifying and counting what needs to be measured.

Supplemental Data Where data are not available in the case management system, they can be collected manually. Three effective manual data collection processes are:

- Requiring attorneys and self-represented parties to file information sheets setting forth data that would not otherwise come to the attention of the court. These sheets should be on forms designed by the court;
- Recording hearing outcome data on court calendars and having a staff member compile the data on a weekly or monthly basis. This method is particularly effective in recording appearance and continuance data; and
- Reviewing court case management information system records or paper court files for individual cases and extracting data from them.

Survey, Focus Group, and Interview Data

Some combination of surveys, focus groups, or interviews is advisable before and after the commencement of the pilot program to obtain the views of:

- Civil case litigants
- Members of the civil bar whose cases were heard during the relevant time period
- Court staff and judges

A standardized approach to survey, focus group, and interview research is being developed as part of this methodology by the IAALS and the NCSC. While it is tempting to think that anyone can develop a good set of questions, this kind of qualitative research is not as simple as it might appear, and it is not inexpensive to administer and analyze in a meaningful way. Designing surveys, interviews, and focus groups requires expertise in order to elicit results that are not biased by the way questions are worded, the order of the questions, and the responses permitted. Without a standardized survey, interview, or focus group protocol for litigants and for the bar, courts will not be able to compare results across jurisdictions.

Surveys, interviews, and focus groups should allow the evaluators to collect information concerning characteristics and demographics of the respondents, their assessment of the effectiveness and fairness of various aspects of civil practice in the court, their assessment of the effectiveness of each newly implemented procedure, and their assessment of the cost impact of new procedures.

Surveys, interviews, and focus groups of lawyers will obtain their assessments of the impact of new rules and procedures on the practice of civil litigation, and the views of civil case litigants on the process, costs, and outcomes of the reformed the civil justice process will provide additional perspective. A combination of paper, telephone and Web-based methods can be used to reach attorneys and litigants and assure broad participation.

Monitoring Cases and Outcomes Participating courts should expect to prepare and review statistical reports on a periodic basis, e.g., monthly, quarterly, or annually. The frequency of reports will depend on what is being monitored and whether corrective action will be taken based on them. Providing too much information can be as problematic as providing too little, distracting the decision maker from that which is truly important. For the purpose of managing cases in the discovery phase, the court will likely want reports on a monthly basis that allow judges and court managers to determine if the revised rules are having their desired effects. For cases proceeding through the court, measures like frequency and number of continuances may be valuable on a monthly basis, while reporting time to disposition may be more appropriately reported on a quarterly basis.

Conclusion

Setting up a pilot project in a way that will allow the objective measurement of the actual outcomes of implementing a new civil rule or civil case management process is not an impossible task. It requires careful planning at the outset, careful gathering of data during the baseline period and the pilot period, and careful analysis of the data gathered. Courts should consider forging a partnership with organizations whose staff has the formal training and appropriate research experience necessary to ensure a successful evaluation. Funding may be available from a variety of federal, state, or private sources to support evaluation projects of this kind.

The problems facing the civil justice system are profound, and the solutions must be well thought-out and tested in advance of broad implementation. When courts evaluate their experiments with new procedures, they are both advancing the cause of their own litigants and bar members and contributing to the body of information that will ultimately serve similar court users across the nation.

UNIVERSITY OF DENVER | INSTITUTE for the ADVANCEMENT of the AMERICAN LEGAL SYSTEM

Institute for the Advancement of the American Legal System
University of Denver
2044 E. Evans Avenue, HRTM Bldg., #307
Denver, CO 80208-2101
Phone: 303.871.6600 www.du.edu/legalinstitute

NCSC National Center for State Courts

National Center for State Courts
300 Newport Avenue
Williamsburg, VA 23185
Phone: 800.616.6164 www.ncsc.org