Chapter 9
Evaluation of Clinical Decision Support

David F. Lobach

Abstract The value and financial impact of clinical decision support systems need to be better understood. This understanding can be formulated through evaluations of decision support systems. This chapter discusses a framework for identifying an evaluation strategy and selecting an evaluation type in order to address important questions concerning clinical decision support systems (CDSS). The chapter also summarizes the existing literature regarding what is known about evaluations conducted on CDSS and about the features of systems that are associated with successful implementations. In order to assist the reader with framing an evaluative study of a CDSS, the chapter includes a multistep approach to formulating an evaluation illustrated by a parallel description of an actual CDSS evaluation study. Finally, the chapter concludes with the identification of several challenges and pitfalls that can be specifically associated with CDSS evaluations.

Keywords Clinical decision support systems • Summative evaluation • Formative evaluation • Randomized controlled trial • Clinical outcomes • Evaluation methodology • Study design

This chapter will explore three aspects of evaluating clinical decision support systems (CDSS): a description of a framework for the evaluation of CDSS; a review of evaluations that have been done on CDSS; and an outline of an approach for conducting an evaluation of a CDSS. The objectives of this chapter are: (1) to provide a context for the value of CDSS evaluation and expose the reader to the types of evaluations that can be performed as well as common outcomes from CDSS evaluations; (2) to summarize findings from CDSS evaluations particularly focusing on systematic reviews and meta-analyses looking at the effectiveness of CDSS and at features that are associated with successful CDSS; and finally, (3) to discuss issues related to performing an evaluation of a CDSS along with an example of a CDSS

D.F. Lobach, M.D., Ph.D., M.S., F.A.C.M.I. (✉)
Klesis Healthcare, Durham, NC, USA

Community and Family Medicine, Duke University Medical Center,
Box 3886 DUMC, Durham, NC 27710, USA
e-mail: david.lobach@klesishealthcare.com

E.S. Berner (ed.), *Clinical Decision Support Systems*, Health Informatics,
DOI 10.1007/978-3-319-31913-1_9

evaluation study. The ultimate goal for this chapter is to help the reader to appreciate the scope of CDSS evaluation, the work that has already been done in evaluating CDSS, and the process of conducting an evaluation of a CDSS.

9.1 Strategies for CDSS Evaluation

From the most simplistic vantage point, CDSS evaluation entails the systematic collection of information about one or more specific aspects of a CDSS in order to gain further insight and understanding about the system. The ultimate intent of CDSS evaluation is to inform or influence decisions related to CDSS design, use, implementation, or effectiveness through the use of empirically collected data. Establishing a framework for CDSS evaluation depends on the question(s) the evaluation seeks to resolve. The process involves identifying a specific evaluation strategy model that defines the intent of the evaluation, and then selecting an appropriate evaluation methodology. Evaluation strategy models can focus on scientific/experimental questions, management issues, qualitative attributes, or user experience [1]. While these strategies are separated for the purposes of discussion, many evaluation projects will draw from multiple strategies in order to gain a full understanding of the impact of a CDSS.

The scientific or experimental models often tend to be most familiar. These models rely heavily on the collection of data that enables systematic comparisons through experimental and quasi-experimental research designs. This model strategy can address issues such as the effectiveness of CDSS to impact decision-making, the effect of CDSS on care process measures, the impact of CDSS on clinical outcomes, or the determination of which system components are most effective.

Management-focused evaluation models seek to understand the impact and effects of a CDSS in the larger context of an overall organization. These contextually-oriented evaluations can generate insights related to workflow disruption, productivity impact, and/or process changes. Findings arising from a management-focused evaluation strategy provide awareness regarding the practical implications of CDSS implementation on work practices. Such findings can be used to refine and modify workflow, to identify bottlenecks or disruptions in service, or highlight newly gained efficiencies.

Qualitative strategies rely on observation and generate data through subjective human interpretation. For example, the qualitative strategy can be used to generate observations pertaining to the effect of CDSS on work processes, decision making, and care delivery. In many instances, qualitative data are systematically reviewed and categorized in order to discover recurring themes and identify trends or patterns.

Participant-oriented strategy models seek to gain understanding regarding the CDSS user experience. These strategies can rely on direct user input in the form of validated survey instruments or on interviews and discussions in order to obtain user feedback. For analysis, user feedback is often categorized in order to find common

themes across multiple users. These studies are exceedingly valuable for refining CDSS so that they will be user-friendly and effective with the intended recipients of the technology.

9.2 Types of CDSS Evaluations

At the highest level, CDSS evaluations can be divided into one of two categories: formative evaluation and summative evaluation [2]. The category selected depends on the ultimate purpose of the evaluation. Formative evaluation refines and improves attributes of the CDSS technology. As implied by its name, formative evaluation seeks to mold and shape a CDSS in most cases prior to its full implementation. Formative studies can follow several different approaches. Needs assessment determines what need or needs the CDSS should address, the extent of the needs, and what approaches could be used to meet the identified needs. Evaluability assessment explores the feasibility of a CDSS evaluation and how such an evaluation can be structured [1]. A structured conceptualization study establishes the most appropriate target population for the CDSS and can begin the process of identifying potential outcomes from system use. An implementation evaluation explores the fidelity of the CDSS in context and serves to identify potential barriers or pitfalls for actual system use. A process evaluation determines the impact of delivering the CDSS on work processes and procedures. These two latter study types are particularly useful for identifying potential problems with a CDSS implementation before they become disruptive in a production environment.

The techniques used for formative evaluations depend in part on the types of questions that are being addressed. Commonly used techniques that could support several of the formative approaches include brainstorming, focus groups, think-aloud sessions, structured interviews, nominal group techniques, concept mapping, and Delphi polls [1]. With regard to techniques for specific approaches beyond the previously identified general methodologies, needs assessment can be accomplished through surveys, interviews with stakeholders, and gap analysis; implementation evaluation may benefit from model creation or simulation studies; and process evaluation may be informed through qualitative and quantitative monitoring from information systems.

Summative evaluations, in contrast to formative evaluations, seek to assess effects or outcomes resulting from a CDSS. Often summative evaluations are the evaluation type that comes to mind when one thinks about evaluating a CDSS. The most familiar summative evaluations focus on the effect of CDSS on process measures or clinical outcomes. Several different study designs can be used to conduct summative outcome evaluations, including randomized controlled trials (parallel group, crossover, or cluster), quasi-experimental investigations (non-randomized trial, historical before – after trial, or time series trial), and observational studies (cohort, case-control, or case series) [3]. Other less familiar types of summative evaluations include impact evaluations, designed to assess intended or unintended

consequences of a CDSS in the production environment; cost-effectiveness or cost-benefit analysis, focusing on the financial consequences of a CDSS implementation; secondary analysis, using existing data to explore new research questions about a CDSS; and meta-analysis, integrating outcome results from multiple studies in order to assess an overall impact from CDSS as a whole.

9.3 Types of Outcomes Assessed in CDSS Evaluations

An important aspect of any CDSS evaluation is the primary metric used to answer the question behind the evaluation. Formative and summative studies reporting evaluations of CDSS have looked at a wide variety of potential outcomes [4]. As Bright et al. indicate, the outcomes identified in published CDSS studies can be grouped into the following seven primary categories:

- Clinical (Length of stay, morbidity, mortality, health-related quality of life, adverse events)
- Health care process (Adoption/implementation of CDS-recommended preventive care/clinical study/treatment, patient adherence to CDSS recommendation, impact on user knowledge)
- Health care provider workload, efficiency, and organization (Number of patients seen/unit time, clinician workload, efficiency)
- Relationship-centered (Patient satisfaction)
- Economic (Cost, cost-effectiveness)
- Health care provider use and implementation (User acceptance, satisfaction, and use and implementation of CDS) [4].

Many studies report on more than one type of outcome, even though one particular outcome may be selected as primary for the purposes of determining the study sample size. Among the 311 CDSS evaluation studies included in a recent review, clinical outcomes were included in 20 % of randomized controlled trials, 35 % of quasi-experimental trials, and 45 % of observational studies. In contrast, healthcare process measures were included in 86 % of randomized controlled trials, 75 % of quasi-experimental studies, and 69 % of observational studies [4].

9.4 Findings from Systematic Reviews of CDSS

Systematic reviews of evaluation studies of CDSS have explored both the effectiveness of CDSS to impact process measures and clinical outcomes and the features of CDSS that are associated with a significant clinical impact. Studies evaluating the effectiveness of CDSS have been reviewed through the Agency for Healthcare Research and Quality-sponsored Evidence-based Practice Center at Duke University [5]. For this review, a CDSS was defined as “any electronic system designed to aid

directly in clinical decision making, in which characteristics of individual patients are used to generate patient-specific assessments, recommendations that are then presented to clinicians for consideration." Examples of CDSS included in the Lobach et al. review included the following: alerts, reminders, order sets, drug-dosage calculations, and care summary dashboards that provide performance feedback on quality indicators or benchmarks [5]. A total of 160 manuscripts representing 148 unique studies were identified and abstracted for inclusion in the systematic analysis. This review determined that CDSS evaluations demonstrate "strong evidence that CDSS… are effective in improving health care process measures across diverse settings using both commercially and locally developed systems" [5]. Relatively little evidence was available to show the impact of CDSS on clinical outcomes and health care costs. The evaluative studies from this review showed that CDSS now have demonstrable effectiveness not just at academic medical centers with locally developed CDSS but also across diverse community healthcare settings using commercially developed CDSS tools [5].

In addition to the AHRQ-sponsored review noted above, CDSS evaluation studies have been serially reviewed by Haynes and colleagues [6–8]. These reviews have demonstrated the historical evolution of CDSS and their impact. The initial review from 1994, a systematic analysis of CDSS evaluation studies that included only 28 controlled trials of CDSS evaluations showed that CDSS improved clinician performance through preventive care, reminders and computer-aided quality assurance [6]. Only three of the ten studies that assessed patient outcomes showed significant improvements. The next systematic review of CDSS evaluation studies was published in 1998 [7]. This review included 68 controlled trials. The use of CDSS was found to have beneficial effects on clinician performance in 66 % (43 out of 65) of the studies. Benefits were shown for drug dosing decision support in 9 of 15 studies, diagnostic assistance in 1 of 5 studies, preventive care reminders in 14 of 19 studies, and on general medical care issues in 19 of 26 studies. Benefit of CDSS was found in 6 of 14 studies that assessed patient outcomes. Of the negative eight outcome studies, Hunt et al. determined that only three had sufficient power greater than 80 % to detect a clinically important effect [7]. The next serial review was published in 2005 [8]. This review of CDSS evaluation studies encompassed 100 randomized and nonrandomized trials, including 97 controlled trials assessing clinician performance. In 64 % of the studies, CDSS improved clinician performance for diagnosis, preventive care, disease management, drug dosing, or drug prescribing. Of the 52 trials that assessed patient outcomes, only 7 reported improved outcomes with CDSS and no reports showed benefit on major outcomes such as mortality. Most of the studies focusing on clinical outcomes were insufficiently powered to detect a clinically important difference [8].

Most recently, evaluation studies focusing on CDSS associated with electronic health record systems (EHRs) have been systematically reviewed by Moja et al. [9]. This review identified 28 randomized controlled trials in which rule-or algorithm-based CDSS were integrated with EHRs and outcome measures reflecting morbidity, mortality, or economic impact were assessed. Sixteen trials measured mortality rates, but found no statistically significant impact. In the nine trials that assessed

morbidity, defined by the authors as occurrence of illness (such as pneumonia, myocardial infarction, or stroke), progression of diseases, or hospitalizations, a small but statistically significant effect was detected. In the 17 trials that reported economic outcomes, differences in costs and health service utilization were detected, but the magnitude of the effect was small. Across all economic outcomes, Moja et al. concluded that there was no consistent advantage for EHRs with CDSS compared to those without CDSS [9].

In summary, evaluations of the impact of CDSS to date have shown that these systems can improve care process measures using both locally developed or commercial CDSS in academic and community-based care settings. More research is needed to determine the effect of CDSS on clinical outcomes and on the economic consequences of using these systems.

With regard to CDSS features that contribute to success, an initial systematic review of CDSS evaluation studies reporting on 71 comparisons of control versus CDSS and focusing on features of CDSS that are associated with a clinical impact was published by Kawamoto et al. [10]. This study showed that CDSS including five specific features were more likely to improve clinical practice than systems without these features. Kawamoto et al. found that the features associated with an impact on clinical practice included the following: "(1) automatic provision of decision support as part of clinician workflow, (2) provision of decision support at the time and location of decision making, (3) provision of recommendations rather than just an assessment, (4) use of a computer to generate decision support, and (5) provision of decision support as part of clinician workflow" [10].

CDSS evaluation studies that identified features associated with successful CDSS implementations were also included in the AHRQ-sponsored review of CDSS evaluation studies [5]. This report confirmed the effectiveness of three features previously identified by Kawamoto et al. namely automatic provision of decision support as part of clinician workflow, provision of decision support at the time and location of decision making, and provision of a recommendation and not just an assessment. The Lobach et al. review also identified six additional features that correlated with successful CDSS systems. These features included:

- Integration with charting or order entry system to support workflow integration,
- No need for additional clinician data entry,
- Promotion of action rather than inaction,
- Justification of decision support via provision of research evidence,
- Local user involvement in the development process, and
- Provision of decision support results to patients as well as providers [5].

Lobach et al. noted that many of the evaluation studies included CDSS with more than one feature, and thus it was difficult to determine the importance or impact of any one individual feature [5].

A more extensive systematic review reporting on 148 CDSS evaluation studies was published by Roshanov et al. [11] This review found that CDSS were associated with success when (1) the system was developed by the authors of the study, (2) the system provided advice to both patients and clinicians, and (3) the system required

a reason for overriding advice. Roshanov et al. also found that "advice presented in electronic charting or order entry systems showed a strong negative association with success" [11]. In contrast to the findings reported by Kawamoto et al. [10], Roshanov et al. found that neither advice automatically provided in workflow nor advice provided at the time of care were associated with successful CDSS [11].

To date, efforts to identify the most critical features associated with successful CDSS have been inconclusive in that the largest studies have reported conflicting results. The findings do imply that features that integrate CDSS into existing workflows and justify the CDSS recommendations lead to greater success; however, more evaluations deliberately focusing on the impact of specific CDSS features are needed.

9.5 Approach to Conducting an Evaluation of a CDSS

There is no "one-size fits all" approach for evaluating CDSS since each system has nuances about its operation and use, and every implementation environment has unique differences and challenges related to evaluation. Accordingly, the following section does not provide specifics for formulating a CDSS evaluation, but instead offers seven guiding questions to facilitate planning an evaluation that accommodates the implementation environment. These guiding questions have been used over the years to assist informatics graduate students with developing thesis and dissertation projects as well as for the formulation of plans for several successful research grants for CDSS development and evaluation. One of the studies resulting from a research grant funded by the Agency for Healthcare Quality and Research will be used as an illustrative example [12]. This project assessed a CDSS that was in use for population health management among 20,000 Medicaid beneficiaries in Durham County, North Carolina. This project was possible in part because every Medicaid enrollee was included in a care management network so that every patient had an assigned medical home and an assigned care manager. Care managers were allocated to a specific population by clinic site, patient age (i.e., pediatric or adult), and native language (i.e., English or Spanish). The decision support tool was designed to detect missing care services such as regular hemoglobin A1c assays for patients with diabetes and potentially inappropriate use of healthcare resources such as emergency department visits for low severity indications or ambulatory care-sensitive conditions. The CDSS tool received administrative claims and enrollment data from the North Carolina Medicaid office monthly and daily admission, discharge, and transfer (ADT) data through a Health Level 7 (HL7) interface from the hospitals and emergency departments in the region [13]. When sentinel events, defined in this context as notable activities that deviate from expected or optimal care pathways such as hospitalizations for ambulatory care sensitive conditions, low severity emergency department visits, missing recommended laboratory tests or services, were detected by the system, notifications for each index patient were communicated to the appropriate care manager via secure email. Sentinel events were

prioritized and filtered so that each care manager would receive only the top 20 most important events each day for his/her patient population.

Question #1: *What will be the impact if the CDSS is successful*? Since every CDSS is designed with an intended purpose, this first question seeks to define what the anticipated result would be if the CDSS worked optimally. After "success" is defined, the corollary question is whether or not this success can be measured directly. If direct measurement is not possible, what surrogate measures could be used to reflect the desired success? In the example study, the goal of the CDSS was to increase the completion rates of recommended care services, and decrease potentially inappropriate emergency department (ED) use and hospitalizations [12]. In addition, success would also result in decreased overall medical costs because expensive ED visits and hospitalizations would be converted to primary care visits [12].

Question #2: *What data are needed to show/measure success or the surrogate outcome*? In order to determine whether or not success has been achieved, empirical data need to be collected and analyzed to assess whether or not the CDSS fulfilled its intended purpose. The data for measuring success need to be defined and characterized. As a reference, the types of outcomes that have been measured in published CDSS studies were summarized above in section 9.3. For the example study, the primary measure was to be emergency department utilization rates. In addition, we also included secondary measures such as low severity ED rates, hospitalization rates, the completion of appropriate medical services, and medical costs in total and broken down across categories of emergency department services and ambulatory care [12].

Question #3: *How can these data be obtained*? One issue to decide is what the ideal dataset would be for measuring a desired outcome; however, obtaining the needed data for all potential study subjects can be challenging. It is important that the evaluation data be available from a common source across both intervention and control subjects. In many instances, supplemental data are collected for the intervention group as part of the research protocol. These data, however, cannot be used to determine the primary study outcome since they are not available for the control group. If the data needed to measure the primary outcome are not available, a different primary outcome or a surrogate for the primary outcome may need to be selected. For the example study, emergency department utilization rates could be readily calculated from Medicaid claims data as could hospitalization rates, completion of medical services, and even medical costs. These data were available for both intervention and control subjects in the target study population [12].

Question #4: *What type of evaluation is possible in the environment of the CDSS*? Now that the desired outcome for measuring the impact of the CDSS has been defined and the necessary data for calculating this metric have been identified, several pivotal issues need to be addressed that determine the type of study that can be performed in the CDSS environment. As discussed above, the likely study type would be a summative evaluation in order to quantitate the impact of the CDSS on the primary outcome. Within the summative evaluation type, a researcher needs to select an appropriate study design. As a general rule, the most rigorous study design

possible is desirable. The next step is to define precisely who the study participants should be. For CDSS evaluation studies, this step can be challenging in that the CDSS content is often delivered to a clinician, but the desired impact is assessed on patients. Whichever study subject group is selected, this group needs to serve as the unit of randomization if a randomized controlled trial design is selected. For the example study, a randomized controlled trial was selected as the study design since the research team had complete control over the distribution of the CDSS recommendations [12]. We opted to use patients as our unit of randomization since the data from the study outcomes would be based on data pertaining to individual patients (i.e., ED utilization rates). As is common with randomized controlled trials involving CDSS evaluations, we then needed to ascertain whether significant contamination would occur since the individuals receiving the CDSS notifications, the care managers, could have contact with patients assigned to both the control and intervention groups. The issue of contamination is addressed below in Question #6.

Question #5: *How should the outcome data from the study groups be compared*? This question seeks to determine what type of data analysis will be necessary in order to ascertain whether or not the impact of the CDSS intervention was significant. The answers to this question will define the analytical approach for comparing the primary outcome measure between study groups. These answers will also dictate what the study sample size and duration need to be in order to achieve an appropriate level of statistical power for the primary outcome. The sample size and study duration are ultimately dictated by the event rate of the primary outcome measure. At this juncture, the researcher needs to determine what a clinically significant difference in the primary measure would be between the intervention and control groups. A sufficient number of primary outcome events needs to occur in order to provide an opportunity to establish a statistically significant difference. For the example project, [12] we were able to determine that the rate of ED utilization within the population was 42.7 visits per month per 1,000 Medicaid beneficiaries. From preliminary studies, we concluded that a change in emergency utilization rates of 20 % would be feasible and would be clinically significant. We estimated that we would be able to enroll approximately 80 % of the available study subjects in the trial. With this number of subjects, we determined that the study duration would need to be 9 months and the power to detect a significant difference would exceed 80 % for an odds ratio for event reduction of 0.80 based on a two-sided test of proportions with a significance level of 0.05. For making comparisons between study groups, estimates for intervention impact on study outcomes were based on generalized estimating equations with a working correlation matrix to account for clustering within families [12].

Question #6: *How can contamination be avoided or controlled*? As mentioned above, under Question #4, the challenge of contamination is prevalent in many randomized controlled trials of CDSS since the patients are often the unit of randomization, but the CDSS intervention is conveyed to clinicians. Clinicians often have contact with patients who have been randomized to the intervention as well as with patients randomized to the control arm. In some instances, it is possible to randomize clinicians and assign all of the clinician's patients to a particular study arm. The

challenge with this approach is that patients are sometimes seen by different clinicians and can then crossover between study groups, leading to a new type of contamination. A cluster randomized trial design in which entire clinician groups or clinic sites are randomized as a unit can control for crossover contamination; however, this approach can significantly limit sample size because the sample size lies somewhere between the number of clusters randomized and the number of patients for whom outcome measures will be determined. In the example study, [12] the contamination question was whether or not a care manager would have contact with subjects randomized to the intervention group and subjects randomized to the control group, and if so, whether this contact would have significant effects on the study outcomes. We determined that the care managers could have contact with both intervention and control subjects; however, this contact with both groups was felt to invoke minimal contamination since the recommendations for intervention were highly patient specific (e.g., pediatric patient X had a low severity emergency department encounter for fever yesterday) and care managers would not know to initiate contact with control subjects without receiving the CDSS guidance. Of note, we did make one adjustment in our unit of randomization such that the randomization was actually based on family units since receipt of an intervention for one family member could potentially impact health behaviors of other family members [12].

Question #7: *What is the economic impact or return on investment* (*ROI*) *from the CDSS*? With the increasing emphasis on controlling healthcare costs, inclusion of measures that will reflect the economic impact of a CDSS intervention are becoming increasingly important [14]. In general, the economic measures can be tied to the changes anticipated in the primary outcome. Basically, if the desired impact is achieved, what are the cost implications? For the sample study, [12] we were seeking to decrease emergency department utilization rates. A decrease in emergency department utilization should contribute to significant cost savings. For this study, we opted to look at total costs for patients in each study arm recognizing that saving costs in one area could lead to increased costs in a different area. For our study, we hypothesized that decreased emergency department utilization rates could lead to increased visits at primary care clinics. In addition, other hypothetical cost implications resulting from our CDSS interventions could be increased rates of testing and potentially increased pharmaceutical costs because intervention subjects would be likely to have more contact with the healthcare system [12].

As mentioned above, these seven questions are intended to assist with the formulation of a CDSS evaluation study. They are not prescriptive with regard to how an evaluation should be conducted for a given CDSS and system environment. Table 9.1 summarizes the seven evaluation questions.

As the "rest of the story" regarding the sample CDSS evaluation project, we did not show a decrease in total emergency department utilization rates or costs across the entire population as our primary study outcomes. However, when we drilled down and looked at the impact of the intervention on low severity ED utilization

Table 9.1 Questions to guide the formulation of a CDSS evaluation

1. Define result if system is successful
Can this success be measured directly?
If not, what surrogate measure can be used to reflect success?
2. What data are needed to show/measure success or surrogate outcome?
3. Where can these data be obtained?
4. What type of evaluation is possible in system environment?
Define study design (historical control, randomized controlled trial)
Identify study participants
5. How should these data be compared?
Define analytical approach
Estimate sample size
Determine statistical power
6. How can contamination be avoided or controlled?
7. What is the economic impact/ROI?

rates, we found a statistically significant reduction in the intervention group relative to the control subjects (8.1 vs. 10.6/100 enrollees, $p<0.001$). In further sub-analysis focusing specifically on the pediatric population, we discovered a statistically significant reduction in total ED utilization rates (18.3 vs. 23.5/100 enrollees, $p<0.001$), translating to a reduction in utilization rate from 28 % to 23 % in the intervention arm. We also demonstrated a statistically significant total cost savings of $500 per pediatric intervention subject during the 9-month study. When the cost saving results were extrapolated across the entire pediatric population, we determined that the annual cost savings for Medicaid would be the on the order of $1.5 million in a single county [12].

9.6 Challenges Associated with Evaluation of CDSS

While evaluation is critical for developing, implementing, assessing impact, and establishing ROI for CDSS, conducting studies of CDSS is fraught with unique challenges. To begin with, performing a rigorous study of CDSS with a simultaneous control arm is often difficult because CDSS tools tend to be components of larger systems. In many instances, it is difficult to allow access to CDSS for some users and not for others. This challenge has increased further with the integration of CDSS capabilities into EHRs. Rarely can an individual site or investigator control the access to CDSS resources within a commercial EHR product thus limiting opportunities for a simultaneous control arm. In almost no instances can the intervention subject be blinded to the intervention. As a consequence, evaluators of

CDSS systems often need to settle for historical controls with a before-after clinical trial study design.

As a second challenge, even when the CDSS can be provided to one group of users, but not another, contamination becomes a problem as illustrated above. As already described, the direct user of CDSS advice is frequently the clinician, even though the impact of the CDSS is assessed on patient outcomes. In clinical settings, it is not unusual for patients to receive care from several clinicians within a practice and for clinicians to care for patients of their colleagues. As a result, non-intervention patients can be exposed to clinicians receiving a CDSS intervention and intervention clinicians can have encounters with patients from both the intervention and control arms of the study.

A further consequence of the discordance between the clinician as the recipient of the CDSS intervention and the patient as the unit of randomization and focal point for the outcome measures is the challenges that arise for data analysis. Special analytic techniques are frequently needed in order to unravel the potential dependencies and co-variation that can arise in these settings. Experienced statisticians are needed on the team in order to navigate the analytic quagmire.

An additional difficulty is that definitive randomized controlled trials are often large studies and can be quite expensive and difficult to organize [15]. As a consequence, many evaluations of CDSS rely on less rigorous study designs and thus tend not to supply the best evidence for CDSS impact.

Another challenge with CDSS evaluation is ensuring adequate exposure to the intervention. In many systems, clinical decision support is an adjunct to the primary workflow and can be easily overlooked or neglected by busy clinicians. In order to have a definitive study of a CDSS, high use rates of the CDSS are required. Over the years, many studies have failed to show impact because the "dose" of CDSS that was "consumed" by the target clinician study subjects was too low. Most of these studies are assessed with an intention-to-treat framework and show no impact for the overall effect. When sub-analyses are conducted and the results for substantive CDSS users are isolated, a positive impact is often detected.

A sixth challenge for evaluating CDSS also relates to problems concerning CDSS usage. In many studies the targeted outcome is a relatively infrequent event such as an adverse drug event for a specific pharmaceutical. In order to have sufficient instances in which the CDSS intervention is activated, the CDSS needs to be regularly used by clinicians so that the CDSS is engaged when the infrequent event occurs [15].

A final and often unanticipated challenge is receiving institutional review board (IRB) approval for CDSS evaluation studies, especially randomized controlled trials. Many CDSS evaluation studies are conducted using computerized information systems in clinical settings that may serve large populations of patients. At the outset, it is often unknown which patients will prompt a CDSS-driven recommendation. In many instances, all patients serviced by the information system need to be included in the study population and thus randomized in the event a CDSS recommendation is generated (or, could be generated, but is withheld for control subjects) for one of these patients. It becomes unrealistic and impractical to consent all of the

Table 9.2 Special challenges arising from evaluation studies on CDSS

Difficulty conducting randomized controlled trials due to limited capacity to selectively turn the system on and off for subsets of users
Contamination in randomized controlled trials
Discordance between CDSS target users (e.g., clinicians) and the target for CDSS impact measurement (e.g., patients)
RCTs of CDSSs are expensive and difficult to conduct
CDSS sometimes have limited use by clinicians mitigating their impact
Need for widespread use of CDSS to generate sufficient power especially for relatively infrequent event outcomes such as adverse drug events
Obtaining Institutional Review Board approval

patients who are receiving care through the information system associated with the CDSS. In these scenarios, IRB review boards may find it difficult to allow a study to be performed on subjects who have been randomized into a clinical trial without the subjects' consent. In such instances, a waiver of consent will need to be justified in order to conduct the study. Such waivers can be justified because these studies promote an evidence-based standard of care and involve minimal risk to the study subjects. As an illustration of this challenge, the example study described above [12] required 9 months in order to obtain Institutional Review Board approval. Many conversations and meetings were necessary to help the IRB understand the nature of a technology clinical trial. Ultimately, a waiver of consent was granted for the trial. Table 9.2 summarizes the challenges for evaluation of CDSS.

9.7 Conclusions/Observations

The rapid expansion of biomedical knowledge and the desire to utilize this knowledge expeditiously at the bedside necessitate the use of tools to support decision making in the delivery of healthcare. Evaluation of these decision supporting systems is becoming increasingly important in order to demonstrate the impact these systems have on the health care triple aim of increased quality, improved health, and lower costs [16]. CDSS evaluations also provide insights regarding how to make systems more effective and efficient for seamless integration into the care delivery process.

Multiple strategies exist for pursuing evaluation of CDSS depending on the particular issues that need to be explored. Strategies can focus on scientific or experimental questions, management concerns, qualitative system attributes, or user experience. After an evaluation strategy has been identified, the type of evaluation to be conducted needs to be determined. In general, evaluations can be formative, focusing on the development and refinement of the CDSS, or summative, exploring the impact of the system on selected outcomes. The types of outcomes that have been reported in the biomedical literature include clinical parameters, healthcare

process measures, workload and efficiency metrics, stakeholder satisfaction, economic impact, and implementation factors.

Published evaluations of CDSS have shown that these tools can improve clinician performance of care processes such as ordering appropriate tests, making correct diagnoses, selecting effective medications, and complying with preventive service recommendations. Initially these benefits were only seen with locally developed systems in academic medical centers. The research evidence now demonstrates that commercial systems can also achieve these benefits and that the benefits can be obtained in community-based settings. The available data on clinical outcomes such as morbidity and mortality is less compelling. A few studies have demonstrated reduced morbidity and decreased length of stay with the use of decision support. Some studies have also shown economic benefits from the use of decision support, but in most instances, the benefits have been relatively small.

Direct evaluation of CDSS has led to the identification of system features that are associated with a positive impact. While the findings from these studies are somewhat inconsistent across studies, there is evidence to suggest that providing decision support for both clinicians and patients, integrating decision support into the clinical workflow, and requiring justification when decision support recommendations are not followed, are all associated with system success.

Conducting an evaluation of a decision support system requires identification of the most salient impact expected from the system, identifying a source of data from which to measure this impact, and designing a study that will generate these data empirically. CDSS evaluations present some unusual challenges in that the systems are often used by clinicians, but the system impact is assessed on patients. This dissonance between the system user and the unit for measuring impact has ramifications for study design, sample size calculations, subject randomization, data analysis, and controlling contamination. In addition, conducting randomized controlled trials to assess decision support is challenging because creating a study environment in which a simultaneous control can be available is often impossible since systems are used across entire enterprises and cannot be selectively turned on and off for a subset of users or patients.

In spite of the challenges associated with evaluating CDSS, these evaluations are critical for the advancement of the field. Multiple types of approaches can be used for evaluating CDSS. A growing body of evidence is becoming available that shows the value of these systems for impacting care processes. More evaluative studies are needed to assess the impact of CDSS on clinical outcomes and economic measures. In addition, evaluations are needed to determine how to create tools that can more efficiently and effectively be integrated seamlessly into the clinical workflow.

References

1. Trochim WMK. Research methods knowledge base: introduction to evaluation. 2006. http://www.socialresearchmethods.net/kb/intreval.php. Accessed 03 July 2015.

2. Spaulding DT. Program evaluation in practice: core concepts and examples for discussion and analysis. San Francisco: Wiley; 2014.
3. Mitchell K. Study design and statistical analysis: a practical guide for clinician. New York: Cambridge University Press; 2006. p. 11–32.
4. Bright TJ, Wong A, Dhurjati R, Bristow E, Bastian L, Coeytaux RR, et al. Impact of clinical decision support systems: a systematic review. Ann Intern Med. 2012;157:29–43.
5. Lobach D, Sanders GD, Bright TJ, Wong A, Dhurjati R, Bristow E, et al. Enabling health care decisionmaking through clinical decision support and knowledge management. Rockville (MD): Agency for Healthcare Research and Quality (US); 2012 Apr. (Evidence Report/ Technology Assessments, No. 203.) Available from: http://www.ncbi.nlm.nih.gov/books/NBK97318/. Accessed 15 Sept 2015.
6. Johnston ME, Langton KB, Haynes RB, Mathieu A. Effects of computer-based clinical decision support systems on clinician performance and patient outcome. A critical appraisal of research. Ann Intern Med. 1994;120(2):135–42.
7. Hunt DL, Haynes RB, Hanna SE, Smith K. Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review. JAMA. 1998;280(15):1339–46.
8. Garg AX, Adhikari NK, McDonald H, Rosas-Arellano MP, Devereaux PJ, Beyene J, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. JAMA. 2005;293(10):1223–38.
9. Moja L, Kwag KH, Lytras T, Bertizzolo L, Brandt L, Pecoraro V, et al. Effectiveness of computerized decision support systems linked to electronic health records: a systematic review and meta-analysis. Am J Public Health. 2014;104(12):e12–22.
10. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of randomized controlled trials to identify system features critical to success. BMJ. 2005;330:765–8.
11. Roshanov PS, Fernandes N, Wilczynski JM, Hemens BJ, You JJ, Handler SM, et al. Features of effective computerised clinical decision support systems: meta-regression of 162 randomised trials. BMJ. 2013;346:f657.
12. Lobach DF, Kawamoto K, Anstrom KJ, Silvey GM, Willis JM, Johnson FS, et al. A randomized trial of population-based clinical decision support to manage health and resource use for Medicaid beneficiaries. J Med Syst. 2013;37:9922.
13. Lobach DF, Kawamoto K, Anstrom KJ, Kooy KR, Eisenstein EL, Silvey GA, et al. Proactive population health management in the context of a regional health information exchange using standards-based decision support. AMIA Annu Symp Proc. 2007; 473–7.
14. Eisenstein EL, Anstrom KJ, Macri JM, Crosslin DR, Johnson FS, Kawamoto K, Lobach DF. Assessing the potential economic value of health information technology interventions in a community-based health network. AMIA Annu Symp Proc. 2005;221–5.
15. Berner ES. Clinical decision support systems: state of the art. AHRQ Publication No. 09-0069-EF. Rockville: Agency for Healthcare Research and Quality (US); 2009 June. Available from https://www.healthit.ahrq.gov/sites/default/files/docs/page/09-0069-EF_1.pdf. Accessed 15 Sept 2015.
16. Berwick DM, Nolan TW, Whittington. The triple aim: care, health, and cost. Health Aff. 2008;27:759–69.