Department of Veterans Affairs
Health Services Research & Development Service | Evidence-based Synthesis Program

QUERI

Electronic Health Record-based Interventions for Reducing Inappropriate Imaging in the Clinical Setting: A Systematic Review of the Evidence

January 2015

Prepared for:
Department of Veterans Affairs
Veterans Health Administration
Quality Enhancement Research Initiative
Health Services Research & Development Service
Washington, DC 20420

Prepared by:
Evidence-based Synthesis Program (ESP) Center
West Los Angeles VA Medical Center
Los Angeles, CA
Paul G. Shekelle, MD, PhD, Director

Investigators:

Principal Investigator:
Paul G. Shekelle, MD, PhD

Contributing Investigators:
Caroline Lubick Goldzweig, MD, MS
Greg Orshansky, MD
Neil M. Paige, MD, MSHS
Brett A. Ewing, MS

Research Associates:
Isomi M. Miake-Lye, BA
Jessica M. Beroes, BS

PREFACE

Quality Enhancement Research Initiative's (QUERI) Evidence-based Synthesis Program (ESP) was established to provide timely and accurate syntheses of targeted healthcare topics of particular importance to Veterans Affairs (VA) clinicians, managers and policymakers as they work to improve the health and healthcare of Veterans. The ESP disseminates these reports throughout the VA, and some evidence syntheses inform the clinical guidelines of large professional organizations.

QUERI provides funding for four ESP Centers and each Center has an active university affiliation. The ESP Centers generate evidence syntheses on important clinical practice topics, and these reports help:

- develop clinical policies informed by evidence;
- guide the implementation of effective services to improve patient outcomes and to support VA clinical practice guidelines and performance measures; and
- set the direction for future research to address gaps in clinical knowledge.

In 2009, the ESP Coordinating Center was created to expand the capacity of HSR&D Central Office and the four ESP sites by developing and maintaining program processes. In addition, the Center established a Steering Committee comprised of QUERI field-based investigators, VA Patient Care Services, Office of Quality and Performance, and Veterans Integrated Service Networks (VISN) Clinical Management Officers. The Steering Committee provides program oversight, guides strategic planning, coordinates dissemination activities, and develops collaborations with VA leadership to identify new ESP topics of importance to Veterans and the VA healthcare system.

Comments on this evidence report are welcome and can be sent to Nicole Floyd, ESP Coordinating Center Program Manager, at Nicole.Floyd@va.gov.

Recommended citation: Goldzweig C L, Orshansky G, Paige N M, Ewing B A, Miake-Lye I M, Beroes J M, Shekelle P G. Electronic Health Record-Based Interventions for Reducing Inappropriate Imaging in the Clinical Setting: A Systematic Review of the Evidence. VA ESP Project #05-226; 2014.

This report is based on research conducted by the Evidence-based Synthesis Program (ESP) Center located at the **West Los Angeles VA Medical Center, Los Angeles, CA,** funded by the Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development, Quality Enhancement Research Initiative. The findings and conclusions in this document are those of the author(s) who are responsible for its contents; the findings and conclusions do not necessarily represent the views of the Department of Veterans Affairs or the United States government. Therefore, no statement in this article should be construed as an official position of the Department of Veterans Affairs. No investigators have any affiliations or financial involvement (*eg*, employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties) that conflict with material presented in the report.

EXECUTIVE SUMMARY

INTRODUCTION

There are widespread concerns within Veterans Affairs (VA) healthcare and in non-VA US healthcare that the costs of healthcare are rising at unsustainable rates. One driver of cost is the increasing use of radiology imaging procedures, particularly advanced imaging techniques such as computed tomography (CT) scanning, magnetic resonance imaging (MRI), and CT angiography. Most authorities agree that more appropriate use of certain imaging tests could both improve quality and save costs.

The recognition that more appropriate use of imaging could improve quality and reduce costs has led to the development of interventions to encourage more appropriate radiology utilization. Some of these interventions have made use of the clinical decision support capabilities of electronic health records (EHR). VA has been a leader in the use of electronic health records and clinical decision support.

In recognition of the risks and costs associated with inappropriate imaging, VA leadership has requested an evidence synthesis which evaluates studied methods for reducing inappropriate imaging that center around the electronic health record (EHR). The final key questions are:

Key Question 1: What is the effectiveness of EHR-based interventions in reducing unnecessary or inappropriate imaging?

Key Question 2: Do EHR-based interventions vary in results by system (type of EHR intervention)?

Key Question 3: What are the harms or potential harms associated with EHR-based interventions used to reduce inappropriate imaging?

METHODS

Data Sources and Searches

We searched the references of existing broad based health information technology (IT) systematic reviews covering the period 1995-2011, and also performed a search of PubMed and Web of Science from 2011 to 9/10/2014 using terms including “Medical Informatics Applications[Mesh],” “Decision Support Systems, Clinical[Mesh],” “medical records systems, computerized,” “health information technolog*,” “electronic medical record*,” and “radiology department.”

Study Selection

Participants: Adult population. Studies aimed only at children were excluded. Studies with mixed populations were included.

Intervention: EHR-based interventions for reducing imaging for diagnostic purposes (as opposed to screening) considered inappropriate or unnecessary based on clinical guidelines. This meant that studies seeking to increase the use of radiographic imaging like mammography for breast

cancer screening were excluded. Studies of systems running on personal digital assistants were excluded. Studies of web-based interventions or computerized, stand-alone systems that we judged could be easily incorporated into the EHR were included.

Comparator (study design): Usual care.

Outcome: Rates of imaging procedures judged as unnecessary based on existing clinical guidelines. Studies that reported on changes in appropriateness (as opposed to decrease in appropriateness) were also included. Studies that targeted the use of imaging procedures stated as being overused and then reporting only utilization data were included. Utilization outcomes were considered separate from appropriateness outcomes. A table of all outcomes included as "appropriateness" is in Appendix C.

Timing: All times

Setting: Ambulatory, hospital, and emergency department settings.

Data Abstraction and Quality Assessment

Data were extracted by 2 reviewers, and discrepancies were reconciled with the group. Articles had data abstracted on study design, time period, setting, imaging modality, intervention, comparison, sample size, target of intervention, findings, IT design, data entry for intervention, and implementation characteristics. We assessed the quality of studies by their design and the degree to which they reported information about intervention and implementation characteristics.

Data Synthesis and Analysis

We used as the primary outcome the effect of the intervention on the appropriateness outcome. This could have been the increase in appropriate use or the decrease in inappropriate use; studies rarely reported both. As a secondary outcome, we used the effect on utilization. Random effects meta-analyses were conducted using the H-K variance estimator. After collecting data on the interventions, implementations, and settings, but prior to extraction of outcomes data, we developed 4 hypotheses regarding effectiveness of the intervention, one in each category of intervention characteristics, settings, implementation, and target.

RESULTS

Results of Literature Search

From all sources, we retrieved 1,195 titles. From these, we identified 172 titles as being potentially relevant. After reviewing these 172 abstracts, we identified 105 titles for full text review. Of these, we rejected a total of 82 articles, with 10 rejected for their study design (*eg*, a commentary, editorial, review, *etc*), 27 rejected as being about radiology imaging for screening, one rejected for being in a child-aged population, and 39 rejected as not being EHR-based or not about clinical decision support (CDS) or not about imaging.

Of the 23 articles included, 3 were randomized trials, 7 were time series studies, and 13 were pre-post studies. Seven studies collected data prior to the year 2000, 7 studies included data collection within the past 5 years (2009 or later). Ten interventions targeted what was sometimes called "high cost imaging," which usually included CT and MRI and occasionally nuclear

medicine tests as well. Four interventions targeted pulmonary CT angiography, 2 studies targeted chest x-ray, 4 interventions targeted multiple radiologic investigations, and 3 studies had other radiologic targets.

In 5 studies, the intervention consisted of simply the display of information, such as the cost of tests, relevant guidelines, or an appropriateness rating for the requested radiology examination for that indication. Nine studies displayed patient-specific information about whether or not the requested study was consistent with existing guidelines (or something similar). Four studies included what we characterized as a "soft stop," meaning for radiology orders that the CDS rated as inconsistent with guidelines or inappropriate, the provider needed to enter a reason why the CDS advice was being over-ridden. Five studies included a "hard stop," meaning providers were prevented from ordering radiology examinations the CDS classified as inappropriate without getting approval from some external person, usually a radiologist or senior clinician.

Two studies did not present data sufficient to include in our quantitative analysis, one because it did not present comparative data without the intervention and one because the outcome was an aggregate measure of many tests, and data specific to the radiology targets were not presented. Of the remaining 21 studies, 13 reported an appropriateness outcome and 13 reported a utilization outcome; 5 studies reported both.

Summary of Results for Key Question 1 and 2

Thirteen studies contributed to each pooled analysis, one pooled analysis for appropriateness and one pooled analysis for utilization. Four studies contributed data to both. Our primary outcome was the effect on appropriateness. Nine of the 13 studies reported statistically significant benefits of the intervention, 2 reported a benefit that was not statistically significant, and 2 studies reported no effect. The random effects pooled estimate from all 13 studies was an effect size of 0.48 (95% CI: -0.71, -0.25). This equates to a "moderate" sized effect, according to a conventional classification.

Thirteen studies reported utilization outcomes. Six studies reported statistically significant benefits of the intervention, and 7 studies reported essentially no effect. The random effects pooled estimate from all 13 studies was an effect size of 0.13 (95% CI: -0.23, -0.04). This equates to a "small" sized effect, according to a conventional classification.

We explored 4 hypotheses regarding effectiveness, one each for characteristics of the intervention, the setting (integrated care delivery versus other settings), the implementation process (the use of audit and feedback was the only implementation characteristic with sufficient data to support a stratified analysis), and the radiologic target of the intervention. We had insufficient studies to support robust pooled estimates of individual strata or to support multivariable analyses. Nevertheless, some patterns are apparent.

All of the interventions with a "hard stop" reported moderate-to-large effects on appropriateness. Studies using other interventions reported more variable effects or had insufficient numbers to draw conclusions.

The 3 studies conducted in integrated care settings all reported large effects on appropriateness and small-to-moderate effects on utilization. Studies conducted at the US institutions that are

leaders in health IT, and other settings, produced more mixed results, although all 5 of the 6 studies of appropriateness at the health IT leaders reported statistically significant benefits of the intervention.

There were too few studies using audit-and-feedback to draw conclusions and no apparent pattern in studies of interventions at different radiology targets.

Summary of Results for Key Question 3

Four studies reported on harms associated with their interventions. One study, evaluating a decision support tool to reduce unnecessary pre-operative testing, found that with the intervention there was an increase in the percent of pre-operative chest x-rays inappropriately not ordered. Prior to the intervention 1.9% of patients did not get a chest x-ray when indicated, compared to 9.3% after the intervention. The clinical impact of this is not known. Another study of a decision support tool to reduce abdominal kidney, ureter, bladder (KUB) x-rays identified 12 KUB studies out of a total of 255 performed against the advice of the tool where there were positive findings. Of these 12, six KUB studies were felt to have significantly influenced patient outcomes, making it unclear whether following the locally developed guidance could have endangered the patient. The 2 other studies reported on qualitative information from physician surveys which primarily identified lack of interest in using the decision support tools because of time constraints and perceived inefficiencies.

DISCUSSION

Key Findings and Quality of Evidence

Key Question 1 and 2

Summary of Findings and Quality of Evidence

Twenty-one studies provide moderate-quality evidence that EHR-based interventions can reduce inappropriate test ordering by a moderate amount, and reduce overall utilization by a small amount. Low-quality evidence supports that interventions that include a “hard stop,” preventing ordering clinicians from overriding a decision support determination that a test is inappropriate, and implementation in an integrated care delivery setting, are associated with greater effectiveness. Audit-and-feedback may be a useful implementation tool, but data are too sparse to draw conclusions. We judged the quality of evidence regarding appropriateness and utilization as moderate, due to heterogeneity in the results. We judged the quality of evidence regarding the characteristics as low, due to the sparseness of the data and indirect nature of the comparisons That is, these characteristics have not been tested as *a priori* hypotheses for differential effectiveness within the same study.

Key Question 3

Summary of Findings

There are few data on the potential harms of decision support tools to reduce inappropriate radiology test ordering. Future studies should evaluate for harms – particularly investigating whether guidelines when applied in practice provide unanticipated results, or when there are issues related to workflow, efficiency, or provider dissatisfaction that could impact a decision

support tool's effectiveness. For example, in a study of CDS to prevent drug-drug interactions, the use of a "hard stop" intervention – while effective in changing prescribing – resulted in delays in treatment for 4 patients, resulting in preventative stopping of the study by the Institutional Review Board.[1] Another study, excluded from our review because it assessed a pediatric population, surveyed physicians and found that most felt the CDS was "a nuisance" and "not relevant to the complex or high risk patients they had to treat."[2] This highlights the need for assessment of harms and unintended effects in every evaluation.

Quality of Evidence

We judged the quality of evidence for harms as very low, meaning any estimate is uncertain.

Applicability

Only 3 studies were performed in integrated care delivery settings, and only one study was performed in the VA. However, there is evidence suggesting that interventions implemented in integrated care delivery settings may be more effective than in other settings, indicating VA may realize benefits equal to or greater than the average benefit reported here.

Research Gaps/Future Research

We identified the following research gaps:

Direct comparisons are needed of different intervention characteristics. We found suggestive evidence that interventions with a "hard stop" are more effective than other interventions, but to prove this hypothesis requires testing the 2 methods head-to-head in the same study. This should be easy to do, since randomization can occur at the provider level, and consist of the CDS with and without the hard stop.

More research is needed on possible harms. Harms of a CDS system intervention with a "hard stop" have been reported in other clinical situations. An explicit assessment of harms should be incorporated into every study of interventions.

About half of our included studies collected data more than 7 years ago, and information technology and attitudes about the use of information technology change over time, so data from more recent time periods would be helpful.

One study reported differential effectiveness by target, and this should be assessed in future studies.

Like all health IT evaluations, more information about context and implementation is needed.

Conclusions

Computerized decision support integrated with the electronic health record can reduce inappropriate use of diagnostic radiology testing by a moderate amount. The use of a "hard stop" as part of the intervention, the use of audit-and-feedback as part of the implementation, and use in an integrated care delivery setting may all increase effectiveness. There are few data on the potential harms of decision support tools to reduce inappropriate radiology test ordering. Future studies should evaluate for harms.