Development of a Methodology to Evaluate Provider Adherence in a Trial of a Guideline-Based Decision Support System for Hypertension

A.S. Chan, MD[1], R.W. Coleman, MS Pharm[3], S.B. Martins, MD, MSc[3], A. Advani, MD, MPH[1,3], M.A. Musen, MD, PhD[1], B.B. Hoffman, MD[2,3],H.Y. Yiu, BA, [4]Z.H. Xiao, MS[5], M.K. Goldstein, MD, MS[2,3]

[1]Stanford Medical Informatics, Stanford University School of Medicine, Stanford, CA
[2]Department of Medicine, Stanford University School of Medicine, Stanford, CA
[3]VA Palo Alto Health Care System, Palo Alto, CA
[4]Department of Epidemiology, Stanford University, Palo Alto, CA
[5]Department of Mechanical Engineering, Stanford University, Palo Alto, CA

ABSTRACT

Measurement of provider adherence to a guideline-based decision support system (DSS) presents a number of important challenges. Establishing a causal relationship between the DSS and change in concordance requires consideration of both the primary intention of the guideline and different ways providers attempt to satisfy the guideline. During our work with a guideline-based decision support system for hypertension, ATHENA DSS, we document a number of subtle deviations from the strict hypertension guideline recommendations that ultimately demonstrate provider concordance. To handle the volume and complexity of possible treatment choices in a clinical trial of a guideline-based DSS, we created a modular advisory evaluation engine that automates the interpretation of clinician concordance with the DSS on multiple levels. We believe that understanding these difficulties with complex guideline-based decision support systems is crucial to any valid evaluation of adherence.

INTRODUCTION

Measuring adherence to guideline-based decision support systems (DSSs) at first glance appears to be a straightforward task. Guidelines by definition are designed to document what are considered best practices and thus may be used as benchmarks for comparison of providers against an accepted standard of care. The overall aim of implementing decision support is to improve patient outcomes. Evaluation of a DSS on patient outcomes involves a continuum of measures. At the top level, the most stringent and ultimately most important test of a DSS would be its impact on patient outcomes- that is morbidity and mortality. However, decision support is one small factor among many in determining overall patient morbidity and mortality; it would be quite difficult to measure with accuracy of the effect of a DSS on such distant outcomes. Intermediate outcome measures, such as the impact of the DSS on blood pressure control, are more proximate measures of efficacy. Finally, evaluation of provider adherence to a particular recommendation made by the DSS is the most direct indicator of the ability of a DSS to effect clinical decision-making.

Considering adherence at a primary message level
Most clinical guidelines have at their core a primary message – a particular focus in overall management of a disease or clinical syndrome. In the case of hypertension, the primary message is to achieve improved blood measure control, thus reducing its associated morbidity and mortality. One measure of the guideline-based DSS is to examine the effect of the intervention with regards to the primary message. Consider for example, an intervention for hypertension. Using this approach, one would simply measure blood pressure control during the study period as an indicator of provider adherence to the guideline. This is problematic, however, when one realizes there may be several possible influences on hypertension control external to the effects of the DSS.

Considering adherence at the recommendation level
Perhaps the most direct measure of adherence is an evaluation of the provider adherence to particular recommendations made by the DSS. This method also presents some interesting challenges. Several studies indicate that information from the electronic medical record (EMR), which is used by the DSS to generate recommendations, may be flawed or incomplete [1, 2]. Our own lab studies indicate that in the case of CPRS, certain blood pressure readings such as home readings stated by the patient are not included in the EMR [11]. Thus, a partial or full deviation from the guideline may be legitimate if the provider has access to additional information not available to the DSS.

Thus, it is extremely important to consider the intentions of the provider when assessing their choices for treatment. Shahar [3] makes the distinction between of an *outcome intention*- the patient state that the guideline attempts to achieve, and a *process intention* – the care-provider's actions that will be used to achieve this outcome. Critiques of adherence require a better understanding of deviation from a guideline and how a deviation from the guideline may still represent an intention to meet the guideline's primary message[4]. Truly intelligent quality assessment of guideline adherence considers both the guideline author's primary intention and the different ways providers adhere to the guideline [4, 5]

Considering adherence at the visit level One method of assessing the relationship between guideline adherence and recommendations made by a DSS is to focus on the delivery of care on a visit-by-visit level. This provides for a temporal relationship between the recommendations made by the DSS and the provider's actions in the clinical context in which the DSS advisory was displayed. By subsequently tracking what changes occurred in the patient's care plan, one can infer the effect the DSS had on the provider.

Considering the Impact of Multi-step Guideline DSS on Evaluation

Many trials involving clinical decision support to date have involved single step reminder or alert systems. There are few examples of evaluation of large-scale implementations of complex guideline-based decision support. Maviglia *et al* discuss the implementation of complex, multi-step guideline-based decision support for therapy of hyperlipidemia at Brigham and Women's Hospital (BWH). In their initial studies, they documented that 69% of patients with atherosclerotic vascular disease failed to meet one or more of the National Cholesterol Education Program goals. 2,258 reminders for 690 patients were delivered in the first year of their evaluation [6]. The final evaluation of the ability of their DSS to influence guideline adherence has yet to be described. Micieli *et al* described one of the few multi-center trials of a guideline-based DSS. Implemented at four Italian centers, Micieli demonstrated compliance with a clinical guideline for acute ischemic stroke from the American Heart Association decreased mortality by 15% at six months [7]. The sheer volume of data and varied outcomes in multi-step guideline-based DSS add yet additional complexity to the evaluation process.

Introduction to ATHENA DSS

ATHENA DSS (Assessment and Treatment of Hypertension: Evidence-Based Automation Decision Support System) is a guideline-based decision support system for the treatment of hypertension. Based on widely accepted national guidelines for hypertension (JNC 6 and the Veterans Administration (VA)), ATHENA DSS delivers treatment advisories to clinicians at the point of care. ATHENA DSS achieves this via an interface to the VA CPRS system, an EMR in patient care delivery settings nationwide.

The ATHENA DSS consists of two main components: a hypertension knowledge base modeled in Protégé [8] and a guideline interpreter that applies the information in the knowledge base to the clinical information retrieved from the CPRS to create patient-specific recommendations for a patient encounter, on a visit-by-visit basis [9, 10].

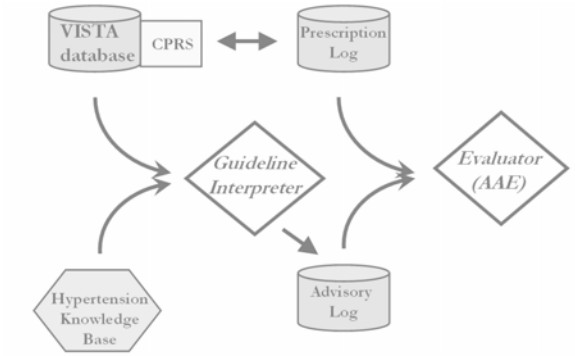

Figure 1: An illustration of the processing of data in the ATHENA DSS. Patient data from the CPRS is considered along with information from the hypertension knowledge base by the Guideline Intrepreter, generating one or more advisories. These advisories are stored in a SQL database and then compared to the prescription log by the Adherence Advisory Evaluator (AAE).

In the ATHENA DSS, we use the concept of a *primary recommendation* to describe the aim of the DSS to deliver a recommendation based on the overall intention of the guideline. If there is a relevant recommendation, ATHENA DSS delivers to the provider a specific list of drug recommendations (the DSS' process intentions) that describes how to achieve the outcome intention of the guideline. No recommendations are made if the patient's care is in complete concordance with the guideline.

ATHENA DSS recommends adding, substituting, or changing the dose of medications in three distinct scenarios: inadequate blood pressure control, choice of therapy that is not concordant with the guideline, and the presence of compelling comorbid conditions that would benefit from the addition of a specific anti-hypertensive medication, without regard to blood pressure control. In situations when blood pressure is well controlled or compelling indications in therapy are not present, the DSS recommends no change in management. These recommendations are presented to the clinicians in the form of a "pop-up" advisory

window superimposed on the EMR that is available for viewing in preparation of a visit or during the patient visit. Both the recommendations of ATHENA DSS and the medication management of the patient after the visit are recorded in a database for subsequent analysis.

ATHENA DSS currently generates advisories for clinicians at multiple clinics within the VA Durham, Palo Alto, and San Francisco Health Care Systems.

GUIDELINE ADHERENCE EVALUATION FOR ATHENA DSS

The primary outcome measure for the ATHENA DSS randomized control trial will be to determine if interaction with the DSS led providers to change the adequacy of blood pressure control and change in overall guideline concordance of the drug therapy.

A subset of the analysis is an evaluation of the intensification of therapy. Intensification of therapy was initially defined as either: an addition of recommended medication, substitution of a recommended medication with another recommended medication, and increase in dosage of a recommended medication.

We were quickly confronted with the tremendous number of advisories that were delivered to providers at our numerous clinical sites. In total, these advisories will be delivered on a daily basis for 12 months at nine dispersed clinical sites. Since each individual patient visit generates a pop-up advisory, the required number of evaluations of provider adherence multiplies rapidly. The complexity of outcomes and sheer volume of data necessitated a solution for an automated approach to guideline adherence evaluation.

Definition of Provider Adherence

The analysis of adherence to ATHENA DSS' three main advisories – add, substitute or increase dosage – requires further discussion. With single step guidelines, the evaluation of adherence is a binary state. The provider is simply evaluated on the presence or absence of appropriate response to the guideline.

We discovered in the course of evaluation of clinician adherence to ATHENA DSS advisories that adherence to the guideline must be defined on multiple levels. In fact, we believe that to truly evaluate the efficacy of the guideline-based DSS, one must also consider the *extent* to which a provider adheres to the guideline. Thus, our analysis includes an evaluation of adherence to the primary recommendation and the extent to which the provider followed any or all of the potentially multiple drug recommendations provided by ATHENA DSS.

1. Addition of a recommended medication
2. Substitution of a recommended medication This includes discontinuation of one drug and replacement with another drug
3. Increase in the dose of the recommended medication

Table 1: Specific advisories provided by ATHENA DSS

Individual assessments of clinician adherence to the add, substitute, or increase dosage messages provided only one level of positive adherence to the guideline. What if, for example, ATHENA DSS recommended that the patient be started on a medication **X**, but the clinician chose to start the patient on a different medication **Y**? Did the clinician simply ignore the recommendation? Or did the clinician recognize that ATHENA DSS recommended intensifying therapy and added a different medication based on patient-specific information not available in the electronic medical record?

Additionally, we know that blood pressure data in CPRS may not be fully updated at the time of advisory generation [11]. Thus, the provider may correctly decide not to alter therapy based on an updated controlled blood pressure value not seen by the ATHENA DSS.

In addition to the three specific messages issues by ATHENA DSS, we have identified a number of scenarios that represent an intention of the clinician to intensify therapy. Table 1 describes the advisories that are explicitly recommended by the ATHENA DSS. Table 2 lists other actions by the provider that suggests that the intent was to adhere to the primary message of the DSS to intensify treatment of hypertension.

1. Addition of any hypertensive medication
2. Increased dosage of any hypertensive medications in the patient's list of medications
3. Partial adherence to the substitution recommendation of a medication, either: a. Discontinuation of the recommended medication b. Addition of the recommended medication as a substitute

Table 2: Additional actions by the provider that indicate a process intention to intensify therapy and thus satisfy the outcome intention of ATHENA DSS.

Evaluation of the criteria presented in Tables 1 and 2 will allow us to calculate both strict adherence to the guideline recommendations and adherence to the primary goal of any hypertension guideline – to intensify hypertensive treatment so as to improve control. Failure to consider these other treatment outcomes would *underestimate* the ability of the DSS to encourage improved clinical treatment.

OUR SOLUTION

To evaluate the multiple potential outcomes of the provider's interaction with ATHENA DSS, we developed the Advisory Adherence Evaluator (AAE), an evaluation engine written in the Java programming language that interprets data generated by the interaction of the hypertension knowledge base, the guideline interpreter, and the VA electronic medical record

ATHENA DSS advisories are initially recorded into a log on a Microsoft Structured Query Language (SQL) Server. This log records patient and visit identifying information, the provider, and the advisories generated by ATHENA DSS. The AAE subsequently retrieves pharmacy prescription data for the patient before and after the visit at which the ATHENA DSS advisory window is displayed. Changes in the prescription data log are then compared with the specific recommendations offered by ATHENA DSS. The AAE subsequently generates a log that measures, on a visit-by-visit level, strict adherence to the ATHENA DSS recommendations and overall intensification of the therapy, even if that action is not explicitly suggested by the ATHENA DSS. This output includes text descriptions of the evaluation and ordinal representations that are easily exported to statistical analysis software such as SPSS.

Total # of Substitution Advisories	1346
Adherence to any portion of the substitution recommendation	432 (32.1%)
Adherence to discontinuation of medication	379 (28.2%)
Adherence to addition of medication	4 (0.3%)
Adherence to both discontinuation and addition recommendations	49 (3.6%)
Nonadherence to substitution recommendation	914 (67.9%)

Table 3: Substitution advisory data from one clinical site demonstrating the effect of considering multiple definitions of adherence

To date, we have tested the accuracy and functionality of the AAE with nine months of data from one of our test sites. The data will remain blinded until the conclusion of the study. Preliminary results demonstrate that the AAE can accurately and efficiently evaluate the multiple potential outcomes from a large volume of provider-ATHENA DSS encounters. We have tested the AAE on a total of 20,369 recommendations involving 8,966 encounters with 66 providers. To illustrate the discovery of additional definitions of guideline adherence, we discovered via the AAE that strict evaluation of a "substitution" advisory from ATHENA DSS (i.e. the discontinuation of a recommended medication and addition of a recommended medication) would have failed to fully describe the number of providers who complied with at least one arm of the DSS recommendation (32.1% vs. 3.6%) (Table 3).

DISCUSSION

We have discovered in our work that to fully appreciate the extent of the provider's adherence to a complex clinical guideline such as the JNC 6 / VA guidelines for hypertension, one must account for the multiple subtleties in treatment selection that a provider considers when selecting appropriate therapy. Any valid measure of adherence must consider the extent to which the provider strictly follows the guideline and the provider's intention to follow the primary message of the guideline. In the case of the JNC 6 and ATHENA DSS, a strict interpretation of the guideline would require a matching of medication choices based on the DSS recommendation. Perhaps more clinically relevant is the provider's intention to follow the guideline's primary message – in this case, the control of hypertension. This is especially important when one considers the body of work that suggests that electronic medical record systems cannot fully capture all of the factors that providers consider when selecting therapy, factors not available to the knowledge base when it generates an advisory. The provider may appropriately and justifiably deviate from the guideline on choice of treatment based on such information.

An example of the complexity of provider adherence evaluation was demonstrated in Table 3. These data reveal important distinctions that are crucial to a valid assessment of provider adherence to a guideline. We have developed and subsequently implemented some of the definitions of guideline adherence that were not initially present at the start of the study. Undoubtedly, in the course of formal data analysis at the end of the trial, we will identify additional factors that will need to be considered in the adherence

analysis. The modularity of the AAE allows for rapid modification without disruption of the core DSS functionality.

As DSS' continue to move from single-step, fairly rudimentary reminder/alerting systems to the complex guideline-based DSS that are exemplified by ATHENA DSS and the BWH work with NCEP cholesterol guidelines, we believe that it will become increasingly important for future evaluations to separate the analysis of strict guideline concordance and analysis of provider treatment behavior. Such analysis may reveal additional information about the true effect of a guideline implementation on clinical outcomes.

We have described the development of a method for evaluating provider adherence to a guideline-based decision support system, ATHENA DSS, in a large multi-center randomized clinical trial. The modularity of the evaluator allows for easy modification of the rules that define adherence to the guideline recommendations without disturbing the underlying structure of the knowledge base and log of patient encounter events. The automation of the evaluation of provider adherence will become increasingly valuable as complex clinical guideline DSS that generate large volumes of complex clinical decision data are implemented.

ACKNOWLEDGEMENTS

This work was supported by grants from the National Library of Medicine LM 05626-02, VA HSR&D CPI 99275 and NIH 5 R01 LM06806.

Views expressed are those of the authors and not necessarily of the Department of Veterans Affairs or other affiliated institutions.

REFERENCES

1. Makoul G, Curry.R, Tang PC., *The Use of Electronic Medical Records Communication Patterns in Outpatient Encounters.* Journal of the American Medical Informatics Association, 2001. **8**: 610-615.
2. Wagner MM, Hogan WR., *The accuracy of medication data in an outpatient electronic medical record.* Journal of the American Medical Informatics Association, 1996. **3**: 234-244.
3. Shahar, Y., *Automated Support to Clinical Guidelines and Care Plans: The Intention-Oriented View*, http://www.openclinical.org/docs/int/briefingpapers/shahar.pdf.
4. Advani A, L.K., Shahar Y. *Intention-Based Critiquing of Guideline-Oriented Medical Care*. in *1998 AMIA Annual Symposium*. 1998. Orlando, FL: C.G. Chute.
5. Advani A, Shahar.Y., Musen MA. *Medical Quality Assessment by Scoring Adherence to Guideline Intentions*. in *2001 Annual AMIA Conference*. 2001. Washington, DC: Hanley and Belfus.
6. Maviglia SM, Zielstorff R., Paterno M, Teich JM, Bates DW, Kuperman GJ, *Automating Complex Guidelines for Chronic Disease: Lessons Learned.* J Am Med Inform Assoc, 2003. **10**: p. 154-165.
7. Micieli G, Cavallini.A., Quaglini S, Guideline Application for Decision Making in Ischemic Stroke (GLADIS) Study Group., *Guideline compliance improves stroke outcome: a preliminary study in 4 districts in the Italian region of Lombardia.* Stroke, 2002. **33**(5): p. 1341-7.
8. Musen MA, Gennari J., Eriksson H, Tu SW, Puerta AR., *Protege II: Computer Support For Development of Intelligent Systems From Libraries of Components.* Medinfo 95, The Eighth World Congress on Medical Informatics, Vancouver, Canada, 1995.
9. Advani A, T., S., O'Connor, M., Coleman, RW, Goldstein, Goldstein, MK and Musen, M.A., *Integrating a Modern Knowledge-Based System Architecture with a Legacy VA Database: The ATHENA and EON Projects at Stanford.* Proceedings of the American Medical Informatics Association, 1999.
10. Goldstein MK, Hoffman B., Coleman RW, Musen MA, Tu SW, Advani A, Shankar R, O'Connor M, *Implementing Clinical Practice Guidelines While Taking Account of Evidence: ATHENA, an Easily Modifiable Decision-Support System for Management of Hypertension in Primary Care.* AMIA Annual Symposium, Los Angeles, CA, 2000: p. 303-304.
11. Goldstein MK, H.B., Coleman R, Gholami, P, Martins S. *Differences Between Initial and Repeat Blood Pressure Measurements During the Same Clinic Visit in Primary Care Practice, 19th Annual VA Health Services Research & Development*. 2001. Washington DC, USA.