# A Systematic Process for Converting Text-based Guidelines into a Linear Algorithm for Electronic Implementation

**David F. Lobach, MD, PhD, MS and Nichole Kerner, MS**
**Division of Medical Informatics, Department of Community and Family Medicine**
**Duke University Medical Center, Durham, North Carolina**

## ABSTRACT

*Clinical practice guidelines are increasingly important for improving the quality and the process of healthcare delivery. Unfortunately, most guidelines are available only in a text-based format, which is difficult to integrate into clinical practice. Computers can facilitate guideline integration into clinical practice; however, this migration to computers requires translating text into intermediary representations. One type of representation that is readily adaptable for computerization is a linear algorithm. This paper describes a systematic process to convert text-based clinical practice guidelines into a linear algorithm with structured content, as an intermediate step to electronic implementation. The process includes: 1) defining applicability criteria, 2) identifying entry points, 3) defining decision points, 4) defining actions, 5) creating a linear algorithm that links decision points and actions, and 6) adding supporting resources. This process has been used successfully to prepare more than two dozen guidelines for computerization. It has been tested by several physicians and informaticians and shown to be transferable to various user groups. The availability of a systematic process to convert text-based guidelines into a structured intermediary format for electronic implementation can facilitate the computerization of guidelines and can inform guideline content developers regarding the critical elements that need to be explicitly stated in guidelines to support electronic implementation.*

## INTRODUCTION

Clinical practice guidelines (CPGs) are becoming increasingly important as a tool to standardize practice patterns and improve the quality and efficiency of health care delivery [1,2]. Unfortunately, most of the available CPGs exist only in a text-based format [3]. In this format, these guidelines are difficult to integrate into the patient care process. Computer-based guideline implementation can facilitate guideline use in practice [2,4]; however, guidelines need to be represented in an electronic format in order to be accessible to a computer [5]. A linear algorithm with structured content is one intermediary process to creating an electronic guideline representation.

Several systems have been developed to implement CPGs electronically [6-11]. Each system has a mechanism for electronically representing guideline content in order to make it computer accessible. Most of these guideline representations utilize a linear algorithm. In addition, efforts to standardize guideline representation through GLIF (The GuideLine Interchange Format) also rely in part on a linear algorithm [12]. Little information is available, however, to direct guideline implementers on how to extract content from text-based guidelines to create a linear algorithm with structured content as an intermediate step to electronic implementation. In this methods paper we describe a systematic process for converting text-based guidelines into a linear algorithm and present results from five of these guideline conversions. This process has been tested by several physicians and informaticians on more than two dozen guidelines.

## METHODS

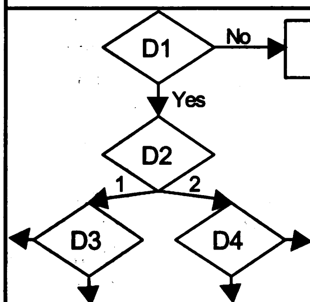
### Systematic Process

A high level diagram of the systematic process for converting text guidelines into a linear algorithm is summarized in Table 1. Initially, the criteria outlining the clinical circumstances and/or defining the appropriate patient characteristics for which the guideline is applicable need to be identified. These *applicability criteria* can include factors leading to inclusion (e.g. cigarette smoker) or exclusion (e.g. pregnancy).

Second, *entry points* into the guideline need to be identified. An entry point is defined as a decision point in the guideline at which a guideline user may initiate interaction with the guideline. The applicability criteria are an entry point common to all guidelines. Examples of other entry points can include follow-up from previous sessions, re-evaluation after additional data collection (physical examination or availability of study results), or various stages of diagnosis or treatment.

Third, *decision points*, discrete expressions in a guideline that form the "logic" of a guideline and determine how the guideline is traversed, need to be identified. Identifying the components of a decision point and converting them into a structured representation is a multi-step sub-process illustrated

*Table 1.* Process for Converting Guidelines into a Linear Algorithm with Structured Content.

| Step | Examples from the Smoking Cessation Guideline |
|---|---|
| **1) Define applicability criteria** | Cigarette Smoker OR Cigar Smoker OR Pipe Smoker |
| **2) Identify entry points** | New effort to quit smoking; Follow-up after recently quitting smoking |
| **3) Define decision points** (see Tables 2 and 3) | Desire to quit smoking is present; Intensive treatment for smoking cessation needed |
| **4) Define actions** | Urge patient to quit smoking in a clear, strong, personalized manner; Arrange follow-up contact within 1 week after quit date; Refer patient to smoking cessation specialist |
| **5) Create a linear algorithm that links decision points and actions** (see Figure 1) |   D1: Applicability Criteria; R1: Guideline not applicable; D2: CASE: Entry points; D3: Desire to quit smoking?; D4: Cessation effort still successful? |
| **6) Add supporting resources** | Sample advice to encourage smoking cessation; Guidelines for prescribing nicotine replacement products |

in Table 2. Each decision component consists of an *element*, a term or phrase that describes a clinical concept; a *critical value*, a term or numeric value to which the element will be compared as a step in traversing the guideline; and an *operator*, a term that defines how the element will be compared to the critical value. Some sample element-critical value pairs (with an operator) are included in Table 2. The clinical concept expressed in an element can be complex. This complexity can be better managed by dividing the element into several specific parts. These parts include a *data parameter*, the most basic concept that still expresses the core meaning of the element, and transformation methods. *Transformation methods* constitute a wide variety of non-temporal information that can be added to a data parameter to "transform" it into the element needed to represent a key concept in a decision component of a guideline. Transformation method types include text modifiers and algebraic expressions. A "modifier" transformation method is included in Table 2. A secondary goal of dissecting elements into data parameters is to link the element to an atomic level variable that would be already stored in a clinical data repository.

Data parameters are further refined by adding

*Table 2.* Converting Decision Points from Text into a Structured Representation for a Linear Algorithm

| Decision point | ***If currently experiencing acute flank pain, and if white blood cell count has been elevated in the past 24 hours, then ...*** | |
|---|---|---|
| Associations | AND | |
| Decision component | Flank pain , acute, current | White blood cell count, elevated, in past 24 hrs |
| Critical Value | present | 12,000 |
| Operator | "=" | ">=" |
| Element | Flank pain , acute | White blood cell count |
| Transformation Method | Modifier - "acute" | (None) |
| Data Parameter | Flank pain | White blood cell count |
| Temporal Constraint | Present, current, (no time) | Present, during past, 24, hours |
| Source | Patient | Reference laboratory |

temporal and source information. Temporal information, referred to in this process as a *temporal constraint*, adds the dimension of time to the data parameter, which in turn is passed to the element to add the dimension of time to the decision components that form decision points in a guideline. The temporal constraint needs to contain four pieces of information: a constraint type (e.g., present or absent), a relativity term (e.g. current), a time interval and the units in which the time interval is measured. Examples of temporal constraints are included in Table 2. The *source information* identifies the anticipated information "creator" from which the data parameter will be collected. Source information may not be relevant for all data parameters. Sample source information is included in Table 2.

In order to complete the definition of a decision point, often several decision components (element-critical value (E-CV) pairs) are combined through pre-defined associations. *Associations* delineate the logical relationships (e.g., "and," "or," etc.) between decision components in a decision point. The common types of associations found in CPGs are listed and illustrated in Table 3. Thus, in addition to defining individual E-CV pairs, the associations between several E-CV pairs need to be explicitly defined in order to represent the intent of each guideline decision point. The combination of E-CV pairs and their associations in essence represent the guideline logic in an individual decision point.

*Table 3.* Types of Associations (logic) Found in CPGs

| ASSOCIATION TYPE | ILLUSTRATION OF TYPE |
|---|---|
| Simple Boolean | A and B and C |
| Complex Boolean | A and (B or C or D) |
| Simple Case | Case 1 A<br>Case 2 B<br>Case 3 C |
| Complex Case | Case 1 A or B<br>Case 2 C or D<br>Case 3 E and F |
| Counted (k of n) | 4 of 7 |

The fourth step in the text-guideline conversion process is to define *actions* that result from resolving the logic expressed by the decision points. The most common types of actions found in clinical practice guidelines are branches, recommendations, orders and referrals. Branches indicate which step(s) follows from the resolution of a specific decision point. This type of action is not represented by our process methodology until the next step, creation of the linear algorithm. Recommendations, orders and referrals constitute the advice that the guideline provides to the user. They all should include information about the target recipient of the advice and the advice content. Recommendations, in addition, need to be specified as clinical, i.e., directly relevant to the intent of the guideline, or procedural, i.e., internal information included to assist the user with traversal of the guideline.

In step 5, the decision points and the actions are integrated by explicitly representing the branch actions in a *linear algorithm*. The algorithm will begin with the applicability criteria represented in the form of a decision point. In most guidelines the applicability criteria will "branch" to a decision point containing a "case" logic statement that lists the guideline entry points. For Boolean-based decision points, branches are labeled as "yes" (i.e., the decision criteria are fulfilled) or "no" (i.e. the decision criteria are not met). The branches arising from "case" decision points are labeled with the number of the "case" that dictates a specific pathway through the guideline. A sample of the beginning portion of the linear algorithm for the smoking cessation guideline is showing in Figure 1 [13].

The sixth and final step of this process is to integrate *supporting resources* into the guideline. Many of the published text-based guidelines are augmented with resources that explain, illustrate or instruct about sections of the guideline. In order to not lose this information when a text-guideline is converted to electronic format, these resources need to be integrated into the linear algorithmic representation of the guideline. An individual resource can be added at one or more points in the linear algorithm as a footnote. The resource footnotes can be associated with decision points or with actions. The inclusion of resources in the smoking cessation guideline is shown in Figure 1.

**Selected Guidelines**

Several diverse CPGs were selected to test this text-based guideline conversion process (Table 4) [13-17]. The selected set of published guidelines addresses both acute and chronic care situations, originate from federal and non-federal sources and have varied levels of complexity and length. Four have been disseminated as published booklets [13-16]. One was published in a peer review journal [17]. All rely primarily on text to convey the guideline content and are not readily amenable to computer implementation as published.

## RESULTS

Our systematic process was successfully applied to more than two dozen diverse guidelines to generate linear algorithms from text-based guidelines. The results of applying this process to five CPGs are summarized in Table 4. In general, the length of the guidelines (determined by number of pages of text) directly correlates with the number of entry points, decision nodes, decision components, and recommendations. In contrast, the number of

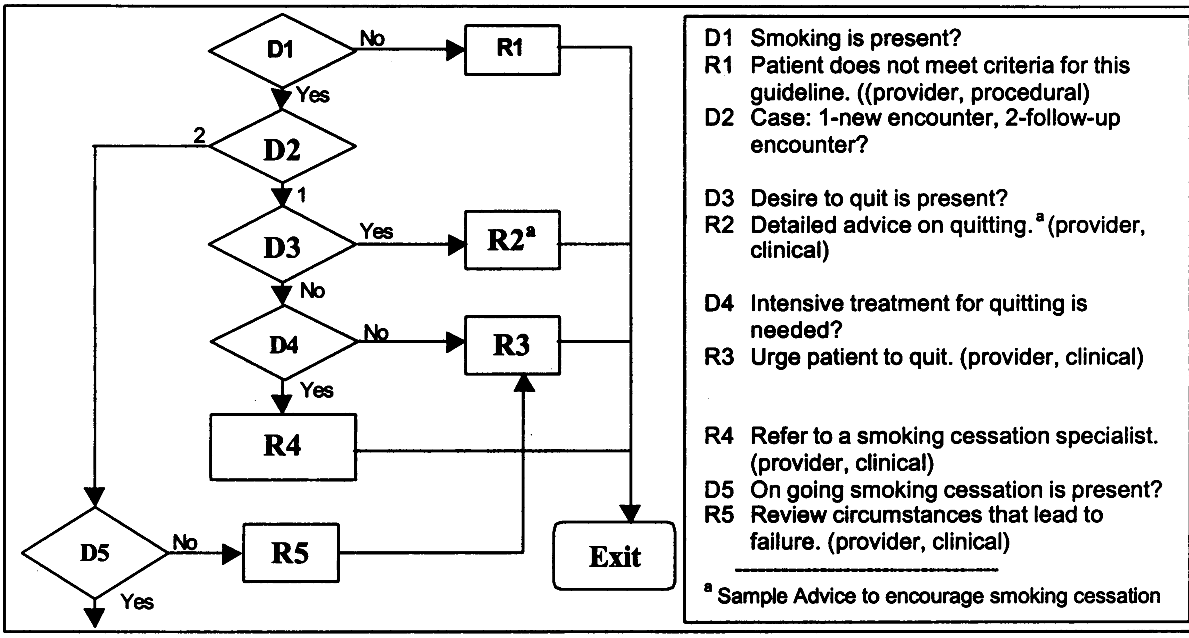



*Figure 1.* Linear Algorithm Derived from Smoking Cessation Guideline.

eligibility criteria and the variety and complexity of the associations (logic) is independent of the guideline length. In addition to applying this methodology to multiple guidelines, it has been used by several physicians and informaticians and shown to be transferable.

## DISCUSSION

In this paper we have described a systematic process to convert text-based clinical practice guidelines into linear algorithms as a pathway to electronic implementation. This process has been successfully used on several guidelines and by several guideline abstractors, which demonstrates that it is generalizable and transferable. While the approach we have outlined has been useful for many guidelines, some guidelines may not be readily amenable to representation as a linear algorithm (and, in turn, may not be amenable to implementation on a computer). Additionally, some published guidelines lack crisp decision points or contain inconsistencies in the guideline logic. In these instances, expert opinion or additional information is needed to convert these guidelines into a computable format.

The resulting representation of a guideline as a linear algorithm can serve as an intermediate stage for several guideline implementation systems. PROforma, PRESTIGE, PRODOGY and EON all reference an algorithmic representation of guidelines as a possible vehicle for entering guideline content into the implementation system [6-9]. Current efforts to standardize guideline representation through the Guideline Interchange Format (GLIF) also reference a linear algorithmic representation of guidelines as a possible input mechanism [12]. Many of the components of the linear algorithm described in this

*Table 4.* Complexity and Length of Five Sample Guidelines

| GUIDELINE | LOW BACK PAIN | ASTHMA | HYPERTENSION | ANKLE | SMOKING CESSATION |
|---|---|---|---|---|---|
| # OF TEXT PAGES | 13* | 106 | 41 | 4 | 30* |
| # OF GRAPHICS | 13 | 56 | 17 | 4 | 41 |
| | | | | | |
| # OF ELIGIBILITY CRITERIA | 2 (AND) | 3 (AND) | 1 | 6 (AND) | 3 (OR) |
| # OF ENTRY POINTS | 9 | 6 | 10 | 1 | 2 |
| # OF DECISION NODES | 35 | 23 | 23 | 3 | 9 |
| # of COMPONENTS | 69 | 93 | 71 | 13 | 13 |
| TYPES OF ASSOCIATIONS (LOGIC) | Simple Boolean<br>Simple Case<br>Complex Case | Simple Boolean<br>Simple Case<br>Complex Boolean<br>Complex Case | Simple Boolean<br>Simple Case<br>Complex Boolean<br>Complex Case | Simple Boolean<br>Complex Boolean | Simple Boolean<br>Simple Case |
| # OF RECOMMENDATIONS | 36 | 35 | 41 | 10 | 23 |

*Equivalent to 8x11 inch page size

paper can be mapped into the other representation schemes. For example, relative to the GLIF2 representation, the "Recommendations" in our model correspond to the "Action Specification" in GLIF2, and the "Associations" in our representation correspond to the "Criterion" in GLIF2. In our SIEGFRIED (System for Interactive Electronic Guidelines with Feedback and Resources for Instructional and Educational Development) tool, we have repeatedly used a linear algorithm as an intermediate step to entering guidelines into our relational guideline representation schema [18].

The systematic process outlined in this paper can serve to minimize the omission of critical information when converting text guidelines into an electronic format. This process can also serve to inform guideline content developers about the critical elements that need to be explicitly stated in their guidelines in order to support electronic implementation [5]. The structure of the linear algorithm and its definitive decision points will also serve to reveal ambiguities that exist in many text-based guidelines.

The proposed process is limited in that it does not necessarily contain all of the information needed to implement a guideline electronically. Nonetheless, it can serve as an initial framework to bridge the gap between text-based and electronic guideline representations. In order to support sharing of guidelines that have been converted into linear algorithms, a universally accepted format and controlled terminology will be needed. Future efforts can build upon our proposed linear algorithm to map the components of our representation to specific components in other functional electronic guideline implementation systems.

**Acknowledgements**

This study was funded in part by a grant from the Agency for Healthcare Quality and Research (R01 HS09436).

**References**

1. Audet AM, Greenfield S, Field M. Medical practice guidelines: current activities and future directions. *Ann Int Med.* 1990; 113:709-14.
2. Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet.* 1993; 342:1317-22.
3. American Medical Association. *Directory of Practice Parameters*. Chicago, IL.: American Medical Association; 1997.
4. Lobach DF, Hammond WE. Computerized decision support based on a clinical practice guideline improves compliance with care standards. *Am J Med.* 1997; 102:89-98.
5. Tierney WM, Overhage JM, Takesue BY, et al: Computerizing guidelines to improve care and patient outcomes: the example of heart failure. *JAMIA.* 1995; 2:316-22.
6. www.rbh.nthames.nhs.uk/PRESTIGE/right.htm. March 6, 2000.
7. Fox J, Johns N, Lyons C, et al. PROforma: A general technology for clinical decision support systems. *Com Meth Prog Biomed.* 1997;54:59-67.
8. Purves IN, Sugden B, Booth N, et al. The PRODIGY project - the iterative development of the release one model. *Proc – AMIA 1999 Annual Symp.* Philadelphia: Hanley & Belfus, Inc.; 1999: 359-63.
9. Musen A, Tu SW, Das AK, et al. EON: a computer-based approach to automation of protocol-directed therapy. *JAMIA.* 1996;3;367-88.
10. Shiffman RN: Representation of clinical practice guidelines in conventional and augmented decision tables. *JAMIA.* 1997; 4:382-93.
11. Zielstorff RD, Teich JM, Paterno MD, et al: P-CAPE: a high-level tool for entering and processing clinical practice guidelines. Partners Computerized Algorithm and Editor. *Proc - AMIA 1998 Annual Symp.* Philadelphia: Hanley & Belfus, Inc.;1998; 478-82.
12. Ohno-Machado L, Gennari JH, Murphy S, et al. The guideline interchange format: a model for representing guidelines. *JAMIA.* 1998; 5:357-72.
13. The Agency for Health Care Policy and Research. Smoking Cessation Clinical Practice Guideline. Number 18; April 1996.
14. The Agency for Health Care Policy and Research Acute Low Back Pain in Adults: Assessment and Treatment Quick Reference Guide for Clinicians. Number 14; December 1994.
15. National Institutes of Health: National Heart, Lung, and Blood Institute. Guidelines for the Diagnosis and Management of Asthma. Expert Panel Report 2; July 1997.
16. National Institutes of Health: National Heart, Lung, and Blood Institute. The Sixth Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. November 1997.
17. Stiell IG, Greenberg GH, McKnight RD, et al. Decision rules for the use of radiography in acute ankle injuries: Refinement and prospective validation. *JAMA.* 1993; 269(9):1127-1132.
18. Lobach DF, Gadd CS, Hales JW: Structuring clinical practice guidelines in a relational database model for decision support on the Internet. *Proc - AMIA 1997 Annual Fall Symp.* Philadelphia: Hanley & Belfus, Inc.; 1997; 158-162.