



Designation: E 2210 – 02

An American National Standard

Standard Specification for Guideline Elements Model (GEM)—Document Model for Clinical Practice Guidelines¹

This standard is issued under the fixed designation E 2210; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This specification covers a document type definition (DTD) that specifies a standard representation for storing and organizing the heterogeneous information contained in clinical practice guidelines. This specification is intended to facilitate translation of natural-language guideline documents into a format that can be processed by computers. It can be used to represent document content throughout the entire guideline life cycle. Information at both high and low levels of abstraction can be accommodated. This specification is based on the guideline elements model (GEM) created at the Yale Center for Medical Informatics and designed to serve as a comprehensive XML-based guideline document representation.

1.2 This specification refers to and makes use of recommendations from the World Wide Web consortium, the W3C.²

1.3 *Standard Guideline DTD*—This specification defines a standard DTD for clinical practice guidelines. The DTD is included in **Annex A1**.

1.4 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory requirements prior to use.*

2. Referenced Documents

2.1 ASTM Standards:

- E 1239 Guide for Description of Reservation/Registration-Admission, Discharge, Transfer (R-ADT) Systems for Automated Patient Care Information Systems³
- E 1384 Guide for Content and Structure of the Electronic Health Record (EHR)³
- E 1633 Specification for Coded Values Used in the Electronic Health Record³
- E 2182 Specification for Clinical XML DTDs in Healthcare³

E 2183 Guide for XML DTD Design, Architecture and Implementation³

2.2 *W3C World Wide Web Consortium:*

XML 1.0 Recommendation⁴

XHTML Basic Recommendation⁵

XLINK⁶

Namespaces Recommendation⁷

XSL/XSLT⁸

Schemas⁹

2.3 *HL7 Standards:*¹⁰

Informative Document: Using XML as an Alternative Message Syntax for HL7, Version 2.3.x
Clinical Document Architecture

3. Terminology

3.1 Definitions:

3.1.1 *document type definition (DTD)*—the formal definition of the elements, structures, and rules for enabling platform-independent data access via XML, or for marking up a given type of SGML document.

3.1.2 *extensible markup language (XML)*—standard from the World Wide Web Consortium (W3C) that provides for tagging of information content within documents, offering a means of representation of content in a format that is both human and machine readable. Through the use of customizable style sheets and schemas, information can be represented in a uniform way, allowing for interchange of both content (data) and format (metadata).

3.1.3 *health level 7 (HL7)*—a standards organization traditionally focused on standards for healthcare information interchange. HL7 messages are the dominant standard for peer-to-peer exchange of clinical text-based information. More recently, HL7 has developed a comprehensive object model of

¹ This specification is under the jurisdiction of ASTM Committee E31 on Healthcare Informatics and is the direct responsibility of Subcommittee E31.35 on Healthcare Data Analysis.

Current edition approved May 10, 2002. Published August 2002.

² <http://www.w3.org>

³ *Annual Book of ASTM Standards*, Vol 14.01.

⁴ <http://www.w3.org/XML/>

⁵ <http://www.w3.org/TR/2000/REC-xhtml-basic-20001219>

⁶ <http://www.w3.org/XML/Linking>

⁷ <http://www.w3.org/TR/REC-xml-names/>

⁸ <http://www.w3.org/Style/XSL/>

⁹ <http://www.w3.org/XML/Schema>

¹⁰ <http://www.HL7.org>

the healthcare enterprise and the first level of an XML clinical document architecture.

3.1.4 *HL7 clinical document architecture (CDA)*—a document markup standard for the structure and semantics of exchanged clinical documents. A clinical document is a documentation of observations and other services with the following characteristics: persistence, stewardship, potential for authentication, wholeness, and human readability. A CDA document is a defined and complete information object that can exist outside of a message and can include text, sounds, and other multimedia content.

3.1.5 *hypertext markup language (HTML)*—the language used in creating a web page. Its origin is an implementation of SGML DTD. It provides tags regarding the way a document should be displayed in the text of an HTML document, which act as commands that a browser interprets when downloading an HTML file.

3.1.6 *namespaces*—provide a simple method for qualifying element and attribute names used in XML documents. This is accomplished by associating a particular tag set by associating a prefix with a URI reference. XML namespaces provides a mechanism for authoring compound documents (documents consisting of elements and attributes from multiple DTDs or schemas) in such a way that will provide global identification without collisions of names that are the same but are used differently.

3.1.7 *parser*—a specialized software program that recognizes markup in a document and differentiates the content from the markup. A parser that reads a DTD and checks and reports on markup errors is a validating XML parser. A parser can be built into an XML editor to prevent incorrect tagging and to check whether a document contains all the required elements.

3.1.8 *schema*—defines the elements that can appear within the document and the attributes that can be associated with an element. It also defines the structure of the document: which elements are child elements of others, the sequence in which the child elements can appear, and the number of child elements. It defines whether an element is empty or can include text. The schema can also define default values for attributes. Schema is a W3C term for the next generation of DTDs.

3.1.9 *stylesheet*—the XSL transformations (XSLT) describes a vocabulary recognized by an XSLT processor to transform information from an organization in the source file into a different organization suitable for continued downstream processing. The extensible stylesheet language (XSL) describes a vocabulary recognized by a rendering agent to reify abstract expressions of format into a particular medium of presentation.

3.1.10 *valid XML document*—a document that is well-formed, with internal or DOCTYPE reference to element definition of tags within the document.

3.1.11 *well-formed XML document*—an XML document that conforms to the syntax as specified by the W3C XML 1.0 recommendation.

3.1.12 *World Wide Web Consortium (W3C)*—develops interoperable technologies (specifications, guidelines, software,

and tools) to lead the Web to its full potential as a forum for information, commerce, communication, and collective understanding.

3.1.13 *XHTML*—HTML documents that are well formed and can be processed by an XML parser.

3.1.14 *XLL/XLINK/XPOINTER*—XLL, the extensible linking language, is divided into two parts, XLinks and XPointers. XLink, the XML linking language, defines how one document links to another document. XPointer, the XML pointer language, defines how individual parts of a document are addressed. XLinks point to a URI (in practice, a URL) that specifies a particular resource. The URL may include an XPointer part that more specifically identifies the desired part or section of the targeted resource or document. XPointer, the XML pointer language, defines an addressing scheme for individual parts of an XML document. XLinks point to a URI (in practice, a URL) that specifies a particular resource. The URI may include an XPointer part that more specifically identifies the desired part or element of the targeted resource or document. XPointers use the same XPath syntax as XSL transformations to identify the parts of the document they point to, along with a few additional pieces.

3.2 Definitions of Terms Specific to This Standard:

3.2.1 *clinical practice guidelines*—systematically developed statements to assist practitioner and patient decisions about appropriate healthcare for specific clinical circumstances.¹¹

3.2.2 *guideline elements model (GEM)*—an XML-based guideline document model that promotes translation of natural language guideline documents into a format that can be processed by computers. Developed at the Yale Center for Medical Informatics, GEM serves as the basis for this specification.¹²

3.2.3 *guidelines interchange format (GLIF)*—a proposed representation for guideline logic created by the INTERMED Collaboratory.¹³

3.2.4 *national guidelines clearinghouse (NGC)*—a website sponsored by the U.S. Agency for Healthcare Quality and Research that disseminates information about qualifying guidelines. It includes a structured vocabulary for describing several aspects of guidelines.¹⁴

3.3 GEM Definitions:

3.3.1 See **Table A1.1** in **Annex A1**.

4. Significance and Use

4.1 *GEM Representation*—The guideline elements model (GEM) was created to unify representations created by health services researchers and by informatics specialists. Specification E 2210 DTD is based on the GEM knowledge representation. It is intended to be:

¹¹ *Guidelines for Clinical Practice: From Development to Use*, Institute of Medicine, National Academy Press, Washington, DC, 1992.

¹² <http://ycmi.med.yale.edu>

¹³ <http://www.glif.org>.

¹⁴ <http://www.guideline.gov>

4.1.1 Comprehensive, that is, capable of expressing all the knowledge contained in a guideline. Existing health services models of guidelines are inadequate for expressing the complexity of knowledge components in sufficient detail to facilitate electronic translations. On the other hand, existing informatics models are insufficient to model constructs that express and support guideline validity. Lack of confidence in the validity of guideline recommendations may ultimately limit end user adherence.

4.1.2 Expressively adequate to express the complexities and nuances of clinical medicine while remaining informationally equivalent to the original guideline. Most tagged elements in the Specification E 2210 DTD store the actual language of the guideline, thereby remaining true to the original. Moreover, this DTD does not require recommendation knowledge to be structured in a temporal sequence, an often artificial transformation necessary for algorithmic representations.

4.1.3 Flexible, that is, a useful model must be able to deal with the variety and complexity of guidelines. The representation should permit modeling at high and low levels of granularity so that guidelines can be interpreted at different levels of abstraction. The Specification E 2210 DTD allows markup using high-level tags or deeper analysis using elements from lower levels in the hierarchy. In addition, the open XML document model can be modified easily if necessary to accommodate missing semantic constructs.

4.1.4 Comprehensible, that is, it should match the stakeholders' normal problem-solving language and allow domain experts to describe their knowledge with little effort. The Specification E 2210 DTD markup does not require knowledge of programming. The markup process parallels physical highlighting of a document and should be learned easily by nonprogrammers.

4.1.5 Shareable across institutions. The use of XML for knowledge representation and markup provides unparalleled cross-platform compatibility.

4.1.6 Reusable across all phases of the guideline life cycle.

4.2 *Conformance*—A document is tested for conformance to this specification by a validating XML parser according to the W3C XML 1.0 recommendation.¹⁵ A conformant document must validate without either well-formedness or validity errors, according to XML 1.0. A conformant document must also conform to constraints expressed within the prose of this specification; however, this specification does not express a formal means of testing conformance to such additional constraints. A document must be valid according to the DTD specified in this specification in order to conform to this specification.

5. Procedure

5.1 *GEM Document Structure*—The architecture of the clinical document is described in another ASTM standard, Guide E 2183. Each clinical document has a header and a body. The structure of the guideline header is specific to the GEM DTD defined in this specification. The structure of the body is unique to each document type. Elements of the body may

contain content or serve as XLinks to external content. Additionally the DTD may make use of namespaces to resolve conflicts with other XML standards and XHTML tags.

5.2 *GEM Architecture*—The root level of this Specification E 2210 DTD is the guideline document. At the next tier, the model defines header and body elements. The <guideline.header> includes identity and developer elements. The <guideline.body> includes purpose, <intended.audience>, <method.of.development>, <target.population>, <knowledge.components>, testing, and revision elements. Each of these elements comprises one or more additional levels of guideline constructs.

5.3 *Elements*—Elements can appear as often as required. Most elements store information that is literally presented in the guideline text itself, for example, release date, name of sponsoring organization, and recommendation text. A small number of metalevel tags provide information about the guideline that has been interpreted, for example, <developer.type>. To distinguish whether an element's content is explicitly stated within the guideline document or was inferred by the person who performed the markup, each element has an attribute called "source." The source attribute can take values of "explicit," "inferred," or in some cases "NGC" (to indicate that the National Guidelines Clearinghouse structured vocabulary is used).

5.4 *Identity*—Information that identifies a particular guideline document and describes it in general terms is clustered within this construct. The identity element includes the guideline's complete title, a citation that references its publication, its <release.date>, its availability (in electronic and print formats) and a person or organization to be contacted for further information. The status element indicates whether the guideline has been updated or revised. Since many current guidelines are released as packages that may include patient education materials, foreign language versions, quick reference guides, and technical reports, a construct for <companion.document> is included. An entry stored in the adaptation element identifies whether the guideline has been adapted from another publication.

5.5 *Developer*—The organization responsible for development of the guideline is identified and described. A <developer.type> element (for example, medical specialty society, federal government agency, managed care organization) provides a structured description of the guideline's sponsor. The formal name of the committee within the developing organization, as well as its members' names and individual or committee expertise, are represented. In addition, sources of financial support for the guideline's development, the names of organizations that have endorsed the guideline, and reference to other organizations' guidelines on the same topic are included.

5.6 *Purpose*—Purpose elements describe the main health practices, services, or technologies addressed by the guideline and the reasons for the guideline's development. Guideline category classifies the major focus of the guideline, for example, diagnosis, treatment, or prevention. The rationale for guideline development (for example, evidence of inappropriate practice or wide practice variation) is subtly different from the

¹⁵ <http://www.w3.org/TR/2000/REC-xml-20001006>

objective of the guideline (for example, to increase use of a particular test, to diminish inappropriate use of a therapy) and either (or both) may be described. The <health.outcomes> element stores the specific health outcomes or performance measures that the guideline is intended to affect. The <available.options> describe the principal alternative preventive, diagnostic, or therapeutic interventions that are available. Exception refers to factors that may permit an exception to be made in applying the guidelines, including home and family situation and constraints on the healthcare delivery system. Strategies, performance measures, and plans for implementing the recommendations may be stored in the <implementation.strategies> element.

5.7 Intended Audience—The <intended.audience> for a guideline refers to the healthcare providers whose behavior the guideline is intended to influence. It includes constructs for both the clinician and the <care.setting> in which a guideline recommendation may be applicable, for example, office, intensive care unit, or a particular health maintenance organization. The <clinical.specialty> element applies the NGC structured vocabulary to categorize the intended users.

5.8 Method of Development:

5.8.1 The validity of a guideline's recommendations is closely tied to concepts incorporated in <method.of.development>. Evidence-based guideline development processes relate recommendations directly to the scientific evidence that supports them. Such constructs are clearly important to developers and implementers, and to end users of guideline recommendations, as they decide whether the recommendations should influence their behavior.

5.8.2 The <description.evidence.collection> element refers to approaches taken by the guideline developers to identify and retrieve scientific evidence. The <method.evidence.collection> stores an NGC structured construct; <number.source.documents> refers to the number of documents identified during evidence collection. The <evidence.time.period> refers to the publication dates of the evidence. The <method.evidence.grading> stores criteria used to gauge the quality of information from different sources and may include a formal <rating.scheme>. The <method.evidence.combination> refers to formal methods of synthesis used to develop summary measures that reflect the strength of scientific evidence, for example, meta-analysis, decision analysis, or formal group judgment techniques.

5.8.3 The <specification.harm.benefit> describes qualitatively the anticipated benefits, potential risks, or adverse consequences associated with implementing the guideline recommendations, while <quantification.harm.benefit> provides an element for storing mathematical models and numeric estimates.

5.8.4 The <role.value.judgment> stores information related to whose values were applied in determining the relative desirability of a health practice. For example, guidelines that optimize healthcare from the point of view of the individual patient, the payor, and society may well differ. Likewise, the specific <role.patient.preference> in determining policies is explicitly declared and stored within this element.

5.9 Target Populations—The <target.population> refers to the group of individuals who are the subject of the guideline recommendations. Eligibility criteria includes <inclusion.criterion> and <exclusion.criterion> that determine the specific portion of the target population for which recommendations are applicable. The NGC specifies gender and age ranges for categorization of the target population.

5.10 Testing—The <external.review> refers to the findings of individuals and groups outside the sponsoring organization that have reviewed recommendations. The <pilot.testing> refers to testing of the guideline's recommendations in clinical settings.

5.11 Revision Plan—The date of “Scheduled Review and Expiration Date” for the guideline recommendations store data that help to determine the validity of the recommendations in light of new evidence.

5.12 Knowledge Components—The <knowledge.components> store and categorize the expert knowledge that is the salient feature of clinical practice guidelines. Knowledge components are classified into three high-level classes, recommendation, definition, and algorithm, because the subelements of each of these call for different approaches to processing. Each of the knowledge components and its subtree in the Specification E 2210 hierarchy is discussed in the following paragraphs:

5.12.1 Recommendations:

5.12.1.1 Recommendations are the unique components that distinguish guidelines from other clinical publications; recommendations are intended to influence practitioners' behavior. When recommendations are analyzed into atomic concepts (and perhaps encoded in a structured vocabulary) they can be executed by a computer's logic.

5.12.1.2 Recommendations can be categorized as conditional or imperative statements. While conditional statements clearly delineate the situations in which they apply, imperatives are broadly applicable to the target population and do not impose constraints on their pertinence.

5.12.1.3 Conditional recommendations can be described in rules that take the form:

If CONDITION then ACTION(S) {because REASON(S)}

5.12.1.4 A condition, in turn, is specified by one or more combinations of a <decision.variable> and its value linked by comparison operators, for example, platelet count less than 50 000. In many cases, the value of a decision variable is not explicitly stated in guideline text but is implied to be true or present.

5.12.1.5 Fulfillment of the condition triggers one or more guideline-specified actions. Reason elements explain why the action has been triggered. The <evidence.quality> that led the guideline developers to call for a particular recommendation and the <recommendation.strength> that they attach to a particular recommendation are tagged in appropriately named elements. The flexibility element describes optional conditions or actions that relate to a particular rule and are often recognizable by the presence of “or” statements in the guideline text. Defining a condition and executing an action often entail an economic burden that can be described in cost elements associated with individual <decision.variables>.