Arden Syntax: Then, now, and in the future

1. The Arden Syntax for Medical Logic Systems

Clinical decision support (CDS) systems [1–4] apply clinical knowledge to patients’ health-related data and generate support of varying extent and depth for the purpose of diagnosis, therapy, prognosis, and patient management. Even administrative and financial decisions may be assisted by such systems if these decisions arise in a clinical context.

Patients’ medical data are usually part of an electronic health record (EHR) or electronic medical record (EMR) system – such as a hospital, laboratory or intensive care information system. Other medical data sources include mobile app systems with smartphone or tablet hardware, biomedical sensors, and medical devices. National or trans-national healthcare systems constitute a gigantic source of medical data of patients and other healthcare clients as well. Today, as opposed to several decades ago, we have giant volumes of patients’ medical data – preferably for further automated processing – in a computable structured form. We now have “big data” in healthcare. Usually, these data, or “big data”, are stored in a repository or data warehouse, technically realized as one or several database systems.

Clinical knowledge to be applied to patients’ medical data also needs a formal representation, a scheme of processing, and a place to be stored. The Arden Syntax for Medical Logic Systems was specifically developed to provide such a clinical knowledge representation and processing scheme. It was developed with constructs that generally define how to access medical data in medical databases, how to process in detail the user-defined clinical knowledge, and how to generate clinical decisions and make them available to the healthcare providers in healthcare situations. Arden Syntax does not define the way that clinical knowledge should be stored technically. However, as data are stored in databases, knowledge is stored in knowledge bases (that might turn out to be databases as well). Modern software technology provides a variety of approaches to knowledge storage and maintenance.

As described later in this special issue, Arden Syntax started at an informatics gathering and academic retreat at Columbia University’s Arden Homestead, NY, U.S.A. – the source of the name Arden Syntax – in 1989; see [5–7] for further reading. Since its inception in 1989, the Arden Syntax has been developed and extended in a series of small and large steps. In 1992, the Arden Syntax version 1.0 was adopted as a standard by the American Society for Testing and Materials (ASTM) as document E1460 – 92, under subcommittee E31.15 Health Knowledge Representation [8].

Two years later, in 1994, Hripcsak [9] published an excellent tutorial on “Writing Arden Syntax Medical Logic Modules.” These medical logic modules (MLMs) act as building blocks for medical knowledge bases. Each module is supposed to contain sufficient knowledge to make at least a single medical decision. This idea was extended later to form so-called medical knowledge packages consisting of interconnected MLMs for more complex CDS [10].

In 1998, sponsorship for Arden Syntax moved to Health Level Seven (HL7) International [11], and this led ultimately to the creation of the Arden Syntax Workgroup [12] in HL7 to supervise this standard. The Arden Syntax Version 2.0 was officially approved by HL7 and the American National Standards Institute (ANSI) one year later, in 1999.

The versions that followed until today, along with their added features are shown in Table 1. The present Arden Syntax version is version 2.10, which was approved by HL7 and ANSI in 2014.

From the beginning, the Arden Syntax for Medical Logic Systems has been intended to be applied in practical clinical settings. To this end, text editors for writing MLMs and interpreters or compilers for parsing and generating computer codes for subsequent processing were written, and the resulting software was integrated into the information technology (IT) ecosystems of medical institutions. Over the years, Arden-Syntax-based software was put into operation at university clinics and also integrated into commercial medical IT systems. Although the number of medical institutions currently using Arden-Syntax-based software for CDS is not known, it may be safely asserted that the Arden Syntax is applied very successfully at the present time. Recent examples of technical realizations of software systems providing Arden Syntax MLM editing and processing are the ArdenSuite IDE and server software by Medexter Healthcare, Vienna, Austria [20,21], the Arden2ByteCode by a group of researchers from the University of Braunschweig and Hannover Medical School, both Germany [22], and the experimental generalization of the Arden Syntax to a common clinical application language by Kraus et al. at University Hospital Erlangen, Germany [23,24].

The aim of this special issue is to recall the beginning and the evolution of Arden Syntax, report on successful clinical applications and lessons that can be learned for further CDS research, and show developments that may be implemented in future versions of Arden Syntax. These objectives are pursued by the following 12 papers, which were prepared for this special issue.

2. Beginning and evolution

In [25], Origins of the Arden Syntax, Hripcsak et al. recall the time when – and why – Arden Syntax was conceived. In the 1980s, at a time of rapid development in knowledge-based and medical expert systems – systems that were procedural and logic-based – researchers recognized the burden of re-creating these knowledge-based systems at every medical institution. The systems provided clinically useful reminders, alerts, and recommendations. They were generally based on clinical rules combined with some medical data access and pre-calculations. The authors explain the starting point: Derived initially from Health
Evolution through Logical Processing (HELP) [26] and the Regenstrief Medical Record System (RMRS) [27], the Arden Syntax defines medical logic that can be encoded as independent rules with the hope of creating a public library. In 1989, an informatics gathering at Columbia University’s Arden Homestead was held and the first version of the syntax defined. At this gathering, the participants developed an ambitious goal: On the one hand, the programming lines should be readable by clinical experts, but on the other hand, also provide powerful computational processing. The kind of processing was derived largely from the programming language APL. The developed syntax was improved and implemented by a number of researchers and vendors in the early 1990s and was initially adopted by the consensus standards organization, the American Society for Testing and Materials (ASTM).

In this issue, Jenders et al. report on the evolution of Arden Syntax and the rationale behind adding new features in Arden Syntax over the years of its life in [28], Evolution of the Arden Syntax: Key Technical Issues from the Standards Development Organization Perspective. The aim is to properly review the design history and the stepwise evolution of the Arden Syntax from the perspective of a standards development organization. Many of the authors participated in this evolution from the beginning or at least over many years. Their perspective is authentic and they report as insiders, not as external observers. The authors report that the initial version of the Arden Syntax for Medical Logic Systems was created to facilitate explicit representation of medical logic so that it could be easily composed and interpreted by clinical experts.

They describe the key goals targeted in developing Arden Syntax. Later, because of the demands of knowledge engineers and programmers to improve functionality related to complex use cases, a variety of features were successively incorporated into Arden Syntax. These changes included expansion of operators to manipulate lists and strings; a formalism for structured output; iteration constructs; user-defined objects and operators to manipulate them; features to support international use and output in different natural languages; additional control features; fuzzy logic formalisms; and mapping of the entire syntax to XML. The history and rationale of this evolution are summarized. In conclusion, the authors write, the Arden Syntax has evolved over time to include a number of powerful features; these make it well suited not only for simple alerts and reminders, but also for more complex CDS. This partly somewhat from the original version of the syntax as being simple alerts and reminders, but also for more complex CDS. This change in perspective increases the utility of this standard for the implementation of CDS. Backwards compatibility has been maintained, allowing continued support of the earlier, simpler decision support models.

### 3. Present state and applications

Since 2004, a large application of Arden Syntax for pediatric-guideline-based care is in use at the Cleveland Clinic, OH, U.S.A. The system is known as Child Health Improvement through Computer Automation (CHICA) and is well described by Anand et al. in [29], Pediatric decision support using adapted Arden Syntax. The authors describe in detail the rationale behind CHICA, its application at the pediatric clinics in their healthcare system, and the extensive use of Arden Syntax MLMs implemented in CHICA. It is argued that today, because of the sheer numbers of pediatric guidelines, and the time and other constraints imposed on the patient’s office visits, guideline-based care is often overlooked or cannot be provided appropriately. In response to this, a computer-based CDS system using Arden Syntax was developed. The approach is as follows: CHICA screens patients together with their families in the waiting room and provides alerts to the physician in the examination room. For this purpose, an adaptation of Arden Syntax to support the production and consumption of patient-specific-tailored documents for every clinical encounter was necessary. These modifications range from new variable definitions, concatenation operator redefinition, to extending prioritization scores and others. The results are impressive: As of the writing of their paper, 429 Arden Syntax MLMs have been written and implemented in CHICA. The system has been used by 755 physicians and other clinical staff over a period of ten years. MLMs – as evidenced by ongoing statistics – were fired more than 11 million times. The authors conclude: “Our results show that the Arden Syntax standard provided us with an effective way to represent pediatric guidelines for use in routine care”.

Schuh et al. report in [30], Clinical decision support systems at the Vienna General Hospital using Arden Syntax: Design, implementation, and integration, on CDS systems at the Vienna General Hospital, Austria. The authors describe the implementation and integration of an Arden-Syntax-based CDS development and processing platform into the newly established hospital information system (HIS) of the Vienna General Hospital, a large tertiary care hospital with about 1800 beds and the main teaching hospital of the Medical University of Vienna, Austria. In the course of a project to replace the former HIS with a new one, it was decided to add CDS functionality to the new system. The project team selected i.s.h.med by Siemens, now belonging to the portfolio of Cerner, as the new HIS, and the ArdenSuite by Medexeter Healthcare as the CDS platform for integration into i.s.h.med. The authors report that, from the beginning, the CDS software was intended to be used for research and teaching as well as for clinical care. As a result of this collaboration between clinicians of the Vienna General Hospital and medical informaticians of the Medical University of Vienna, three CDS applications were developed and implemented using Arden Syntax. They are described in the present paper: (1) Moni-ICU is an automated system for monitoring and surveillance of hospital-acquired infections in the intensive care units (ICUs) for adult patients; (2) TSM-CDS is a system for risk prediction in the formation of cutaneous melanoma metastases; and (3) TacroDS is a CDS system for dosage adjustment with tacrolimus, an immunosuppressive agent given after kidney transplantation. In all three systems, Arden Syntax has been used to model the applied clinical knowledge and carries out the necessary calculations and decision support rules.

From the same institution comes a detailed report by de Bruin et al. [31], Assessing the feasibility of a mobile health-supported clinical decision support system for nutritional triage in oncology outpatients using Arden Syntax. The report examines the feasibility of a mobile health-supported CDS system for nutritional triage in oncology outpatients using Arden Syntax.

### Table 1

Overview of the development of the Arden Syntax since version 2.1.

<table>
<thead>
<tr>
<th>Version</th>
<th>Year</th>
<th>Essential changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>2002</td>
<td>new string operators; reserved word “currenttime” returns the system time</td>
</tr>
<tr>
<td>2.5</td>
<td>2005</td>
<td>object capabilities: create and edit objects; extended markup language (XML) representation of MLMs (except logic, action and data slot)</td>
</tr>
<tr>
<td>2.6</td>
<td>2007</td>
<td>UNICODE encoding; additional resources category to define text resources for specific languages; time-of-day and day-of-week data types; “localized” operator to access texts in specific languages</td>
</tr>
<tr>
<td>2.7</td>
<td>2008</td>
<td>enhanced assignment statement; extended “new” operator to allow easy and flexible object instantiation</td>
</tr>
<tr>
<td>2.8</td>
<td>2012</td>
<td>additional operators for list manipulation; operators to manipulate parts of given date and time values; switch statements; keyword “breakloop” for aborting a loop</td>
</tr>
<tr>
<td>2.9</td>
<td>[13]</td>
<td>2013 fuzzy data types, fuzzy sets, and fuzzy logic; adjustment of all available operators to be able to handle fuzzy data types [14–17]</td>
</tr>
<tr>
<td>2.10</td>
<td>[16]</td>
<td>2014 XML representation of MLMs (including logic, action and data slot) [19]</td>
</tr>
</tbody>
</table>
Syntax. The nutritional monitoring system implements the so-called patient-generated subjective global assessment (PG-SGA) to regularly observe and rate the nutritional state of oncology outpatients. The background is that there is no standard practice in many European hospital wards and outpatient settings to identify early signs of malnutrition in patients with cancer. A smartphone application was developed and used to collect patient-generated data by performing weekly mini-surveys on patients concerning their eating habits. The collected data are combined with clinical data from the Vienna General Hospital’s HIS. Ten Arden Syntax MLMs, each including either a sub-assessment or the final results of the PG-SGA were defined. After processing, the results are routinely shown as a part of clinical care. A three-month pilot clinical trial was performed to test the feasibility of the system. In all, 22 patients completed the clinical study; 91% of the study patients believed that this smartphone application was of high value in detecting cancer-related malnutrition.

An important study using Arden Syntax MLMs was conducted at the University Hospital of Erlangen, Germany, by Castellanos et al. [32], Using Arden Syntax Medical Logic Modules to reduce overutilization of laboratory tests for detection of bacterial infections—Success or failure? Repeated measurements of procalcitonin (PCT) as a biomarker for bacterial infection and especially sepsis is typically used for early detection and follow-up. However, PCT measurements are costly and their overutilization should be avoided. The study describes a CDS system based on Arden Syntax that was developed to determine necessary and also preventable PCT orders. On the basis of the patient’s latest PCT result at a surgical ICU, the time period since this measurement, and the expected PCT trend, the authors assessed how the developed CDS system affects the daily rate of ordered PCT tests. This was done in the course of a prospective study with two on and two off phases. The authors report an 18% reduction of PCT tests within the first four weeks of CDS system use at the investigated ICU. This reduction may have been influenced by raised awareness of the overutilization problem. No reduction of PCT tests was observed during the second on phase. The physician interviews indicated that time-critical ICU situations may prevent extensive reflection about the necessity of each single PCT test.

In [33], Data-driven knowledge acquisition, validation, and transformation into HL7 Arden Syntax, Hussain et al. report on a study aimed at obtaining clinical knowledge from practice data sets for the treatment of head and neck cancer, to validate the acquired knowledge against published treatment guidelines, create logical rules therefrom, and incorporate the rules into a CDS system. The final system is based on Arden Syntax MLMs—also to be shared among different organizations—and is evaluated using real patient data from a HIS. First, a dataset of 1229 patients suffering from oral cancer was used to create a predictive model based on a decision tree algorithm. The resulting model was tested on an independent data set comprising 739 patients. The results of this evaluation went into a so-called refined clinical knowledge model that was then formally converted into four Arden Syntax MLMs. These MLMs form the knowledge core of a CDS system, which was applied to—again—739 real patient cases. Comparison of the individual results obtained with those received by applying each case on the Nauchi-PCT model based on a decision tree algorithm. The result at a surgical ICU, the time period since this measurement, and the expected PCT trend, the authors assessed how the developed CDS system affects the daily rate of ordered PCT tests. This was done in the course of a prospective study with two on and two off phases. The authors report an 18% reduction of PCT tests within the first four weeks of CDS system use at the investigated ICU. This reduction may have been influenced by raised awareness of the overutilization problem. No reduction of PCT tests was observed during the second on phase. The physician interviews indicated that time-critical ICU situations may prevent extensive reflection about the necessity of each single PCT test.

In the next paper by Kraus et al. [36], Using Arden Syntax for the creation of a multi-patient surveillance dashboard, again from the group of researchers at the University Hospital of Erlangen, Germany, the authors demonstrate that the Arden Syntax can in principle—although not primarily defined for this purpose—support multi-patient CDS. Usually, CDS using Arden Syntax MLMs processes the data of individual patients, providing decision support concerning diagnostic or therapeutic healthcare decisions for those specific patients. However, the simultaneous use of decision support functionality on groups of patients is necessary. The authors successfully demonstrated the transformation of Arden Syntax MLMs into other formats that can be used or are demanded by other rule engines. This transformation is technically done by defining and applying XSLT style sheets. The authors anticipate that the development of further XSLT style sheets will support various rule engines.

4. Future development

In [37], Accessing complex patient data from Arden Syntax Medical Logic Modules, researchers Kraus et al. of University Hospital Erlangen, Germany, again explore a new avenue of application for Arden Syntax. The Arden Syntax MLMs are able to retrieve patient data from EMRs to enable decision support. For patient data stored in simple data structures such as lists or tables, the mapping into MLMs is straightforward. For patient data of complex structure such as microbiology
data access, the mapping into MLMs is more difficult. This, the authors write, has triggered the investigation of approaches for providing complex patient data from EMRs inside Arden Syntax MLMs. The authors identified three different approaches to map complex data from electronic patient records onto Arden Syntax variables; each of these approaches successfully mapped a test sample of complex microbiology data. The first approach was implemented in Arden Syntax itself, and the second one inside the interface component of one of the investigated Arden Syntax environments. The third was based on deserialization of XML data. The results show that the most promising approach was to map XML structures onto congruent complex data types of Arden Syntax through deserialization. It may help overcome a major obstacle in integrating CDS functions into clinical information systems. The authors recommend extending the Arden Syntax specification with an appropriate XML data format.

The last of the twelve contributions to this special issue and also the last of three papers in the “Future development” section is a paper by Karadimas et al. [38], User-defined functions in the Arden Syntax: An extension proposal. It suggests an expansion of the present Arden Syntax by user-defined functions. The authors argue that one restriction of the Arden Syntax is that one cannot define functions and subroutines inside an MLM. An MLM can, however, call another MLM, where this MLM will serve as a function. This will add an additional dependency between MLMs. The article explains why the authors believe the Arden Syntax would benefit from a construct for user-defined functions, discusses the need, the benefits and the limitations of such a construct. The result of this investigation was that the authors defined a new grammar that supports user-defined functions. Twenty-two production rules were added to the grammar. A parser was built using the construction of useful parsers (CUP) software tools. Examples are given to illustrate the concepts; all examples were parsed correctly. The authors conclude that it is possible to add user-defined functions to the Arden Syntax in a way that remains consistent with the standard.

5. Conclusions

Although the Arden Syntax is no longer an infant, toddler or adolescent, it still seems young, a bit grown up, and pleasantly possessed of vitality in terms of its development and applications. Over time, the Arden Syntax has become an extended, versatile, and still comprehensible syntax for medical knowledge representation and processing. Various software components have been built around it – from MLM authoring tools, interpreters, compilers, servers to different CDS triggers and patient data access methods. Arden-Syntax-based software has been integrated into commercial health IT and is used in research and teaching such as student and funded projects, medical informatics and eHealth classes, and many others. Despite some open questions and controversies, the Arden Syntax is a very lively structure and continues to evolve. It has been incorporated into modern software environments to become a fully developed, versatile and scalable basis for CDS, and integrated in, interfaced with, and used by health IT in numerous clinical applications. Arden Syntax has proven highly beneficial in real patient care situations and fulfills its original intention of creating a definitional foundation for computerized clinical knowledge to be applied in CDS systems.

An aspect that is largely outside the precincts of the present special issue, but an important extension of Arden Syntax, is known as Fuzzy Arden Syntax. Several references can be found in Table 1, with version 2.9. In one contribution by Schuh et al. [30] in this special issue, Moni-ICU, an automated system for the monitoring and surveillance of hospital-acquired infections in ICUs is described. Moni works with fuzzy sets and fuzzy logic [39,40], but not yet with Fuzzy Arden Syntax. However, two recent publications report on the use of Fuzzy Arden Syntax in creating clinical fuzzy automata [41] and in developing medical fuzzy control systems [42]. There is great potential in the extension of Arden Syntax by fuzzy concepts. First, the linguistic uncertainty due to the unsharpness (fuzziness) of boundaries in medical terms can be formally captured by fuzzy sets. The representation and thus application of these terms becomes more intuitive – let me say even more correct – than sharp, black/white thresholds. Examples are simple medical terms such as fever, leukopenia, leukocytosis, but also more complex concepts such as hypoxemia or clinical signs of a blood stream infection. If, for example, fever is part of a rule triggering the hypothesis sepsis, a clear-cut threshold of greater than or equal to 38 °C will prevent the triggering of this hypothesis when measuring a body temperature of, let’s say, 37.9 °C. This does not correspond to medical thinking and, subsequently, to clinical action. Second, propositional uncertainty, due to given incompleteness in medical conclusions, is captured by fuzzy logic. Thus, medical uncertainty in definitional (not necessarily a contradiction), causal, statistical (not necessarily a probability), and heuristic relationships can be formally described by fuzzy logical relationships and – together with the results of evaluating the defined fuzzy sets with real measured patient data – propagated through the fuzzy logical inference network. Third, even models beyond simple rule systems such as fuzzy automata and fuzzy control systems (see above) can be defined and put into operation using Fuzzy Arden Syntax. This is probably a methodological extension, the founders had never thought of!

Another development that supports the use of Arden Syntax in diverse practical clinical settings is the extension of its interoperability. Interoperability in this context has two aspects: First, it means the triggering of an MLM or a package of MLMs in a CDS situation. If the Arden Syntax software is fully integrated into other health IT software, the trigger – either data-dependent, time-dependent, or by manual action – comes from the host system’s health IT software. If the Arden Syntax software is not integrated, but “only” interfaced, the triggering may come from, for instance, RESTful web services. This immediately brings us to the second aspect of interoperability. Data is generally needed when processing an MLM. The founders of the Arden Syntax used two curly braces {} as syntax structure to indicate that data is needed from outside of the MLM. Expressions inside the two curly braces are usually SQL statements fitting a certain database structure of an EHR or EMR. Today, using new developments in software technology, the needed data may – and this is one data access strategy – be sent directly as a package in response to a RESTful service call.

Nevertheless, the curly braces are still around, actively applied, but also criticized. The need to adapt each data access expression in an MLM – if this MLM is transferred to another place of operation – is a challenge termed the curly braces problem [43–46]. However, over the last 25 years, the inability of health care IT to define and implement a uniform data model that everyone, including Arden-Syntax-based CDS systems, could rely upon has made necessary the curly braces. In this setting, the data translation solution adopted by Arden Syntax must be seen as a virtue. The approach effectively isolates the unavoidable data translation problem from the core medical knowledge providing a workable solution to the challenges of transferring medical knowledge among dissimilar systems.

It may turn out that the so-called curly braces problem will provide path to the next generation of interoperability solutions. Today, the curly braces inside of MLMs allow access to many different data sources – even in parallel. In addition to so-called database connectors that can be called from inside of curly braces, a fast healthcare interoperability resources (FHIR) [47] connector has been added to today’s Arden-Syntax-based CDS software in order to process FHIR resources data in MLMs. Moreover, MLMs can be called from business process model and notation (BPMN) software [48], may call openEHR data warehouses [49], or are called out of CDS Hooks communication exchange processes [50]. The communication always is through modern RESTful web services.

Nevertheless, these improvements in the Arden Syntax over time have not occurred in a vacuum. As sometimes occurs in the world of standards, multiple standards emerge that address similar use cases and
facilitate related functionality. For example, in part because of an increasing emphasis on measuring health care quality in reproducible ways, increasing effort has been undertaken in the development of standards to represent shareable quality measures – so-called eMeasures [51]. After an earlier effort to address this space with the healthcare quality measure format (HQMF), HL7 recently has created the clinical quality language (CQL) [52]. Moreover, regulatory agencies in the U.S.A. have incorporated CQL in their rules and incentive reimbursement schemes for reporting health care quality measurement and instituting CDS. In light of the emerging importance of measuring quality and the interplay between quality improvement and CDS, harmonization may be required between proven solutions such as the Arden Syntax and emerging standards for quality measure representation such as CQL.

However, the Arden Syntax continues to evolve. It is hoped that a uniform patient data model can soon become an intrinsic part of Arden Syntax. Then, MLMs could be freely exchanged between clinical users without fear of data incompatibilities. An early candidate for a uniform patient data model can soon become an intrinsic part of Arden Syntax for medical decisions in XML. Int Med J 2008;77(10):650–6.


-de Bruin JS, Stelzer H, Rappelsberger A, Adlassnig K-P. Creating clinical fuzzy automata with Fuzzy Arden Syntax. AMIA annual symposium proceedings


Nadkarni PM. Metadata-driven software systems in biomedicine – designing systems that can adapt to changing knowledge. London: Springer; 2011.


