01 Mar 2015

An Interoperable System For Automated Diagnosis Of Cardiac Abnormalities From Electrocardiogram Data

Thidarat Tinnakornsrisuphap

Richard E. Billo
Missouri University of Science and Technology, richard.billo@mst.edu

Recommended Citation

T. Tinnakornsrisuphap and R. E. Billo, "An Interoperable System For Automated Diagnosis Of Cardiac Abnormalities From Electrocardiogram Data," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 2, pp. 493 - 500, article no. 6809847, Institute of Electrical and Electronics Engineers, Mar 2015.
The definitive version is available at https://doi.org/10.1109/JBHI.2014.2321515

An Interoperable System for Automated Diagnosis of Cardiac Abnormalities from Electrocardiogram Data

Thidarat Tinnakornsrisuphap and Richard E. Billo

***Abstract*—Electrocardiogram (ECG) data are stored and analyzed in different formats, devices, and computer platforms. As a result, ECG data from different monitoring devices cannot be displayed unless the user has access to the proprietary software of each particular device. This research describes an ontology and encoding for representation of ECG data that allows open exchange and display of ECG data in a web browser. The ontology is based on the Health Level Seven (HL7) medical device communication standard. It integrates ECG waveform data, HL7 standard ECG data descriptions, and cardiac diagnosis rules, providing a capability to both represent ECG waveforms as well as perform automated diagnosis of 37 different cardiac abnormalities. The ECG ontology is encoded in XML, thus allowing ECG data from any digital ECG device that maps to it to be displayed in a general-purpose Internet browser. An experiment was conducted to test the interoperability of the system (ability to openly share ECG data without error in a web browser) and also to assess the accuracy of the diagnosis model. Results showed 100% interoperability using 276 ECG data files and 93% accuracy in diagnosis of abnormal cardiac conditions.**

***Index Terms*—Automated electrocardiogram (ECG) diagnosis, Health Level 7 (HL7), interoperability, medical device communication, ontology.**

I. Introduction

CURRENTLY, 12-lead electrocardiogram (ECG) monitoring systems have proprietary formats for operation and data storage and are manufactured by a multitude of different vendors. The ECG data are recorded, read, and analyzed by different methods depending on computing platforms and software implementation intricacies. ECG output data are not shared among different products, or able to be presented in a ubiquitous manner across heterogeneous computing platforms that do not contain the vendor's ECG software display product.

On a related issue, with respect to disease diagnosis from the output of these ECG systems, automated systems have been developed that can serve as supportive decision aids to physicians diagnosing cardiac conditions. By examining the ECG signal, a number of informative measurements can be derived from the characteristic ECG waveform that, in turn, can lead to a deduction of a specific cardiac condition. Much of this reported research has been limited to development of heuristic models to detect a single specific cardiac disease.

For example, hidden Markov models have been used to examine ECG waveforms to detect abnormal heart rhythm [1]; point scoring has been used for detection of right ventricular hypertrophy [2]; and artificial neural nets and computer interactive redefinition of criteria for myocardial infarction [3], [4]. However, no system has been reported for automated detection of a robust suite of cardiac abnormalities that are recorded through ECG waveforms.

There is a need to share and integrate ECG data among different devices and systems [5], [6], as well as provide a greater level of automated support to physicians that use data from these systems to detect cardiac disease. Such technologies are necessary to advance arising new models of healthcare such as telemedicine and assistive technologies. The availability of such technology provides an overall better quality of healthcare to patients.

II. Objectives

The objectives of this research were to: 1) develop an ontological model and subsequent software system to promote open exchange and presentation of ECG data; and 2) use this ontology as an aid for the automatic diagnosis of a number of common cardiac abnormalities that may be present in ECG output. The general approach was to make use of an emerging standard for electronic exchange of medical data; to develop an ontology for the explicit specification of such a standard; to implement this ontology in a widely accepted machine readable format so that ECG data exchange can be done efficiently without the need of proprietary algorithms or software; and finally, to use the ontology as the basis for automated ECG diagnosis of 37 specific cardiac abnormalities. This research included data from 12-lead ECGs operating both in at-rest situations and in continuous monitoring situations such as ambulatory recordings.

Health Level 7 (HL7) [6] was selected as the medical data exchange standard. The ontology that was developed served to integrate ECG waveform data, HL7 ECG data descriptions, and cardiac diagnosis rules. The ontology was then implemented in the Extensible Markup Language (XML) because of its widespread acceptance for exchanging data on the Internet.

Manuscript received July 16, 2013; revised October 28, 2013, February 3, 2014 and March 6, 2014; accepted April 26, 2014. Date of publication May 2, 2014; date of current version March 2, 2015.

T. Tinnakornsrisuphap was with the School of Mechanical, Industrial and Manufacturing Engineering, Oregon State University, Corvallis, OR 97331 USA (e-mail: thidarat_t@lifetime.oregonstate.edu).

R. E. Billo is with the Office of the Vice President for Research, University of Notre Dame, Notre Dame, IN 46556 USA (e-mail: rbillo@nd.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JBHI.2014.2321515

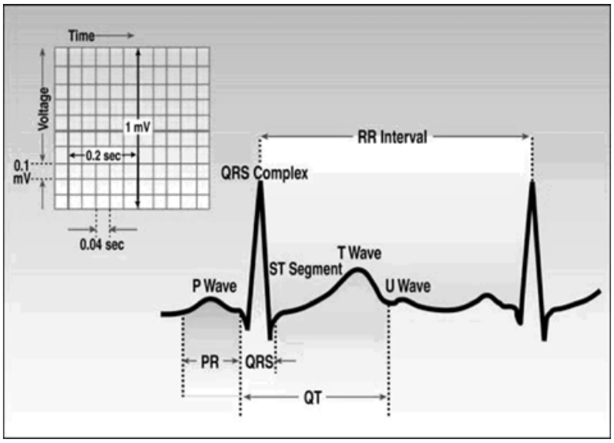


Fig. 1. Standard ECG wave structure (adapted from [8]).

TABLE I
IRREGULAR RHYTHM DIAGNOSIS DESCRIPTION

Abnormal Condition #1:
Irregular Rhythm
General Descriptions
Irregular rhythm is a condition of disturbances in the heart's rhythm.
ECG Measurements
- Duration of the interval between two consecutive QRS complexes
- Duration of the interval between two consecutive P waves of ECG

Criteria
- Irregular rhythm can be determined by accessing whether the RR intervals and PP intervals are regularly spaced.
- If the rhythm is irregular, determine if:
  - o It is occasionally irregular
  - o Regularly irregular (there is a pattern to the irregularity)
  - o Irregularly irregular (there is no pattern to the irregularity)

III. ONTOLOGICAL MODEL FOR REPRESENTATION AND DIAGNOSIS

An ontological model was developed to represent ECG waveform data, explicit measurements of the ECG waveforms, and the diagnostic rules that allow detection of specific cardiac abnormalities. The HL7 point-of-care medical device communication standard [7] was used as the exchange standard. The ontological model served two purposes, allowing both interoperability and diagnosis.

A. ECG Wave Structure

To model ECG output data and then to make inferences as to any cardiac abnormalities that it may represent, one must first understand the structure of the ECG waveform. Fig. 1 illustrates the structure of ECG waves, along with intervals, standard time, and voltage measures [8].

The P wave represents atrial activation. The PR interval is the time from onset of atrial activation to onset of ventricular activation. The QRS complex represents ventricular activation while the QRS duration is the duration of ventricular activation. The *ST*-T-wave represents ventricular repolarization. The QT interval is the duration of ventricular activation and recovery. The U-wave represents the time interval after depolarization in the ventricles.

B. HL7 Medical Device Communication

HL7 provides standards for the exchange and sharing of electronic health information [6]. HL7 was designed to address the interface requirements of an entire health care organization including clinical, financial, and administrative information among heterogeneous computer systems. The standards enable healthcare information interoperability and sharing of electronic clinical and relevant data.

The Lab Automation Special Interest Group (SIG) of HL7 defined a draft set of standards for point-of-care medical device communication, which has concurrently been adopted as a draft IEEE standard [7]. It was intended to provide for open systems communications between medical devices and patient care information systems. The scope of the standard includes nomenclature architecture and a data dictionary for ECG and other clinical areas such as Vital Signs, Respiratory Measurements, and Common Blood Gas Measurements. This research made use of the ECG section of the HL7 point-of-care medical device communication standard that includes both the data dictionary for ECG measurements, as well as enumerations for ECG diagnostics (i.e., abnormal conditions) derived from ECG signals by an ECG monitoring system.

C. Diagnostic Rules

It is important to note that the ECG section of the HL7 point-of-care medical device communication standard makes no causal connection between the ECG measurements and the ECG diagnostics. The research team was required to develop these diagnostic criteria as part of the project. The diagnosis knowledge was collected from available rules obtained from a variety of sources, including cardiology textbooks, interviews with experienced cardiac surgeons, and peer-reviewed papers in the published literature [7]–[14]. An example diagnosis of *Irregular Rhythm* is described in Table I. Diagnostic rules were developed for a total of 37 cardiac abnormalities.

D. Ontology Schema

An ontology-based system representing structure for the presentation, measurement of ECG data, and criteria for diagnosis was developed. The structure of the ontology was synthesized from existing data descriptions as listed in the HL7 medical device communication standard, the technical and clinical literature and recommendations from cardiac surgeons. Fig. 2 depicts the schema for the developed ontology representing the relationships among ECG data and the HL7 standard. The ontology integrates the ECG waveform data with HL7 textual measurement and diagnosis descriptions. In addition, through the type entitled *Criteria*, the developed ontology improves the use of the HL7 medical device communication standard by adding causal linkages between ECG measurements and abnormal conditions that represent a comprehensive set of diagnosis rules that draw upon HL7 measurements.

In the ontology, no preliminary morphological classification of a beat was done. All beats were considered. *Waveform*

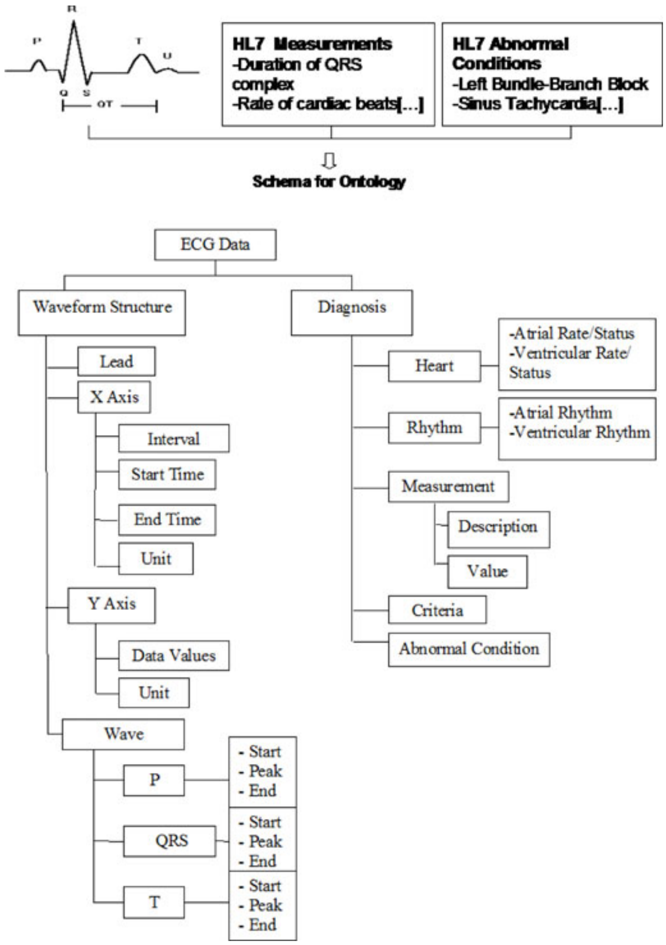


Fig. 2. Schema for ontology.

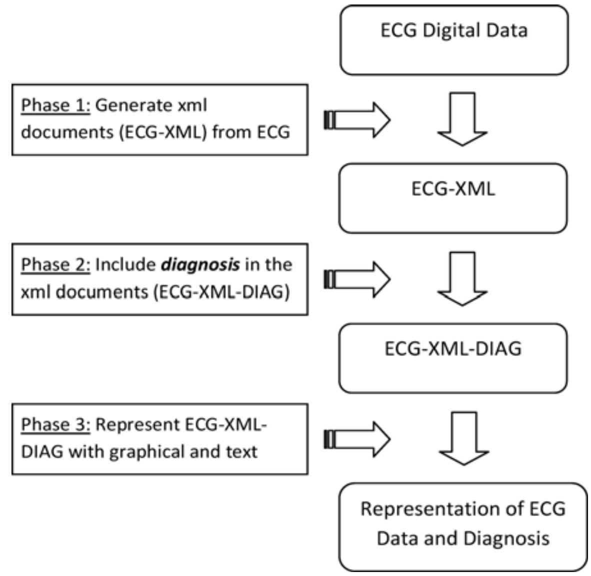


Fig. 3. Software development framework.

TABLE II
ECG DIGITAL DATA

Sample #	Time (second) x-axis	Amplitude (mV) y-axis
1	0.000	-0.270
2	0.003	-0.255
3	0.006	-0.245
:	:	:
3333	20	-0.345

Structure consists of necessary components for ECG plots and associated waves. The structure format was adapted from recommendations in [10]. Lead refers to a vector along which the heart's electrical activity is recorded as a waveform [11], which can be No. 1–No. 12 for each electrode lead. In the type *Wave*, "Start" indicates the sample number that starts the wave; "Peak" indicates the sample number of the peak point of the wave; and "End" indicates the sample number that ends the wave.

The type *Diagnosis* serves to connect HL7 measurements and abnormal conditions through diagnosis criteria. As stated previously, the diagnosis knowledge was collected from available rules obtained from a variety of sources including cardiology textbooks, interviews with experienced cardiac surgeons, and papers in the published literature.

IV. SOFTWARE FRAMEWORK

The software operating process for converting ECG waveform data to a common data file, implementing the ontology, and diagnosing abnormal conditions is illustrated in Fig. 3. Tasks were divided into three phases and are described in the following sections.

A. Phase 1: Generation of XML Documents (ECG-XML) from ECG Data

XML documents were automatically created from ECG x-, y-coordinate data. *XML Schema* [10], [15] was used to restrict and validate ECG data instantiated into XML documents. Waveform data and structure were also encoded in XML Schema. This task made use of the *ecgpuwave* software available from *PhysioNet* [16] to detect the boundary of each wave (start, peak, and end). The XML Schema consists of structures for the components of the waveform drawn from the x and y coordinates of the graphical waveform. The x-axis represents the time interval in seconds while the y-axis represents amplitude in millivolts. Table II shows an example of ECG digital data.

Fig. 4 continues the example by presenting the ECG-XML document showing the instantiation of ECG data above (x, y) and ECG wave data (P-wave, *QRS* complexes, T-wave).

B. Phase 2: Inclusion of Diagnosis in the XML Documents (ECG-XML-DIAG)

In this step, all beats of ECG data in XML format (ECG-XML) were diagnosed by using the diagnosis rules for each abnormal condition listed in the HL7 medical device communication standard. Table III shows examples of ECG measurement properties including duration, amplitude, and appropriate wave intervals. To properly identify abnormal conditions, rate and

```xml
<?xml version="1.0"?>
<ECG xmlns:xsi="http://www.w3.org/2001/XMLSchema-
instance">
  <Lead>MLII</Lead>
  <GraphInfo>
    <XAxis>
      <IntervalScale>0.003</IntervalScale>
      <StartTime>10</StartTime>
      <EndTime>29.997</EndTime>
      <TimeUnit>Second</TimeUnit>
    </XAxis>
    <YAxis>
      <DigitalData>-0.27,-0.255,-0.245,-0.24,
-0.24,-0.245,-0.25,-0.235,-0.225,-0.235,-0.23,
-0.245,-0.245,-0.24,-0.23,-0.225,[..]
      </DigitalData>
      <AmpUnit>mV</AmpUnit>
    </YAxis>
  </GraphInfo>
<WaveComponent>
  <Pwave>
    <Start>14</Start>
    <Peak>29</Peak>
    <End>53</End>
  </Pwave>
  <QRScomplex>
    <Start>70</Start>
    <Peak>89</Peak>
    <End>99</End>
  </QRScomplex>
  <Twave>
    <Start>188</Start>
    <Peak>214</Peak>
    <End>251</End>
  </Twave>
</WaveComponent>
  [..]
</ECG>
```


Fig. 4. ECG-XML document instantiated with ECG coordinate and wave data.

TABLE III
ECG MEASUREMENT PROPERTIES

Waves	Measurements
P wave	• Duration • Amplitude • P-P Interval
QRS complex	• Duration • Amplitude • R-R Interval • P-R Interval
T wave	• Duration • Amplitude • Q-T Interval

rhythm were calculated and included as properties in the type *Diagnosis* as shown in Table IV.

Rules for diagnosis were then added to the system. An example of *ECG-XML-DIAG* is listed in Fig. 5. This example file includes an encoding of the *Irregular Rhythm* description (from Table I) and *Premature Ventricular Contraction* (description not shown).

TABLE IV
RATE AND RHYTHM PROPERTIES

Rate/Rhythm	Measurements
Rate	• Atrial Rate • Ventricular Rate • Status
Rhythm	• Atrial Rhythm • Ventricular Rhythm

C. Phase 3: Representation of ECG-XML-DIAG With Graphical and Text Information

ECG-XML-DIAG documents were converted to output in an Internet browser for presentation purposes. Users are able to see output consisting of an ECG image and appropriate diagnosis without the need of proprietary software. The presentation process was implemented in an Internet browser by using JavaScript, a Java Applet, and PHP Version 4.3.11-1 as programming language tools.

For user's output, both graphical and textual information are presented via a browser. Graphical information consists of an image of the ECG plot along with boundary and status of each wave. Textual information lists heart rates (both atrial and ventricular rates), rhythm status (both atrial and ventricular rhythms), abnormal ECG measurements, and possible findings (normal ECG or abnormal conditions found).

Fig. 6 illustrates graphical output of the process for a single beat of Premature Ventricular Contraction and output of textual information for this ECG data.

V. VALIDATION

The system was validated for interoperability, as well as its ability to correctly diagnose abnormal conditions.

A. Interoperability

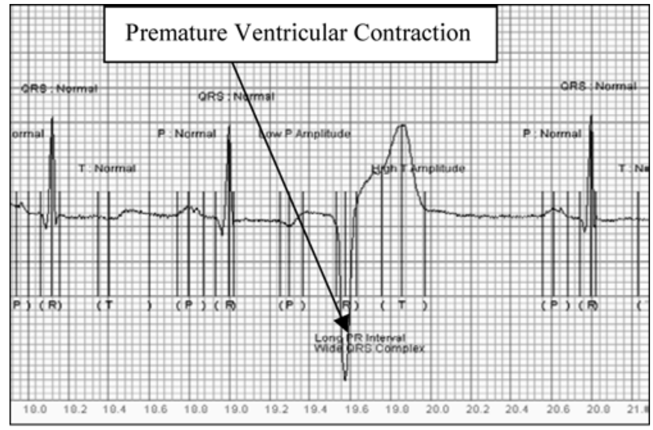
Interoperability was defined as the ability to openly share ECG data in a web browser without the need for proprietary software. With such a definition, users of heterogeneous platforms would be able to use the system via web access without the need for proprietary ECG device software. The researchers' ultimate intention was that, in practice, ECG device manufacturers would provide software drivers for their products that map their ECG device data from their proprietary formats to the standard ontology realized as XML code as described in this research. In this way, the data could be readily exchanged and displayed in any Internet browser. This is common practice in other industries. For example, CAD vendors routinely allow for a user to store or translate design data from their product's proprietary format to an open standard for exchange purposes.

In the present research, ECG waveform data uniquely structured in device files was converted to a common file consistent with the ontology developed from the HL7 device communication standard. The criterion for interoperability was that all ECG data from numerous files went through the system and displayed without error in Internet Explorer.

```
<?xml version="1.0"?>
<ECG    xmlns:xsi="http://www.w3.org/2001/XMLSchema-
instance">
    <Lead>MLII</Lead>
    <GraphInfo>
        <XAxis>
          <IntervalScale>0.003</IntervalScale>
          <StartTime>10</StartTime>
          <EndTime>29.997</EndTime>
          <TimeUnit>Second</TimeUnit>
       </XAxis>
       <YAxis>
          <DigitalData>-0.27,-0.255,-0.245,-0.24,
-0.24,-0.245,-0.25,-0.235,-0.225,-0.235,-0.23,
-0.245,-0.245,-0.24,-0.23,-0.225,-0.22,-0.215,
-0.205,-0.18,-0.155,-0.145,-0.145,-0.145,0.165,[..]
          </DigitalData>
          <AmpUnit>mV</AmpUnit>
       </YAxis>
    </GraphInfo>
<WaveComponent>
  <Pwave>
    [..]
    <Duration>0.13</Duration>
    <Amplitude>0.16</Amplitude>
    <PPInterval>0.87</PPInterval>
    <Measurements>Normal</Measurements>
  </Pwave>
  <QRScomplex>
          [..]

  </QRScomplex>
  <Twave>
     [..]
  </Twave>
</WaveComponent>
        [..]
<Rhythm>
  <AtrialRhythm>Irregular</AtrialRhythm>
  <VentricularRhythm>Irregular</VentricularRhythm>
</Rhythm>
<HeartRate>
  <Atrial>
    <AtrRate>46-120</AtrRate>
    <AtrStatus>Various Rates</AtrStatus>
  </Atrial>
  <Ventricular>
    <VenRate>49-104</VenRate>
    <VenStatus>Various Rates</VenStatus>
  </Ventricular>
</HeartRate>
<Abnormality>Irregular Rhythm,Premature Ventricular
Contraction</Abnormality>
```

Fig. 5. ECG-XML-DIAG document.



(a)

Lead: MLII

<u>Summary of Findings</u>

Heart Rates

Atrial Rate: 46-120 bpm **Status:** Various Rates
Ventricular Rate: 49-104 bpm **Status:** Various Rates

Rhythm

Atrial Rhythm: Irregular
Ventricular Rhythm: Irregular

Abnormal ECG Measurements

- Low P Amplitude
- Long PR Interval
- Wide QRS Complex
- High T Amplitude

Possible Findings

- Irregular Rhythm
- Premature Ventricular Contraction

(b)

Fig. 6 (a) Graphical and (b) textual output.

Interoperability was tested with ECG data from eight databases representing eight different vendor devices that were available in *PhysioBank* [17]. The system converts data from *PhysioBank* because such data was stored in the *PhysioBank* format and did not conform to any open standard. In turn, *PhysioBank* databases themselves represent an archival copy of ECG data contributed by authors of numerous published articles. Each database remains unchanged from the original archived copies provided to *PhysioBank*. *PhysioNet* software provides access to the unique structures of the various databases so as to make the data available to users of *PhysioBank*; however, the data is not interchangeable, thus cannot be viewed without *PhysioNet* software or functions from *PhysioNet* to allow access. So, although our system literally only performs a single conversion of heterogeneous data from the *PhysioBank* system to an open standard (thus demonstrating the feasibility for device manufacturers to develop such drivers for their unique products), if successful, the interoperability test would indicate that it could provide transparent viewing in web browsers, without the need for ECG device vendors' software.

We used *PhysioBank* data of ECG's because they were well characterized, meaning the data themselves were carefully reviewed by physicians to correct any errors so they could be used for assessment purposes. In total, these eight databases provided 276 ECG data files that served as input to the system. The databases were:

1) MIT-BIH Arrhythmia;
2) European ST-T;
3) Long-Term ST;
4) MIT-BIH Noise Stress Test;
5) Creighton University Ventricular Tachyarrhythmia;
6) MIT-BIH Atrial Fibrillation;
7) MIT-BIH Supra-ventricular Arrhythmia; and
8) Normal Sinus Rhythm.

ECG data were a combination of data from patients at-rest and ambulatory with length varying from 10 s to 24 h. Further details on conditions under which ECG data were collected for each of the databases can be found in [17].

B. Diagnosis

This task involved validating accuracy of the system in correctly detecting abnormal cardiac conditions as well as normal ECGs. Discrete data analysis [18] was used to determine the accuracy of the model in diagnosing normal and abnormal conditions. For accuracy, the list of abnormal conditions found by the software for each ECG was compared with their actual known diagnoses that were provided with the ECG data from *PhysioBank.* (The databases contained information about normal ECG or abnormal conditions accurately diagnosed by physicians.)

Accuracy of the model was validated in terms of its sensitivity, specificity, and overall accuracy. Sensitivity is the ability to predict positive results correctly when abnormal conditions exist. Specificity is the ability to predict negative results when the data are actually normal. Overall accuracy is the calculation resulting from combining results from both sensitivity and specificity tests. It represents the percentage of total correct results predicted by the model.

This research used an acceptable error rate of 10% determined from information about ECG decision support systems and algorithms for classifying cardiac diseases in the literature [19], [20] and from recommendations from three experienced cardiac physicians.

Assuming a normal distribution, a sample size that guaranteed the following two probability requirements was determined from the acceptable error rate [21]:

1) P[accept the model when error rate $\geq p_2] \leq \beta$(accept bad lot);

2) P[reject the model when error rate $\leq p_1] \leq \alpha$ (reject good lot).

The first inequality states that the sample size should be large enough so that if the error rate is greater than p_2, there is a very small chance (less than β) of not detecting it. The second inequality states that if the error rate is satisfactory (e.g., less than p_1), there is a low probability (less than α) of rejecting the model. Normally, p_1 ranges between 1% and 5% while α and β range between 5% and 10%. A sampling plan with a minimum sample size of 117 was selected to assure the two probability requirements with 95% confidence intervals were met.

Inferences on population probabilities of discrete random variables were considered. The cell probability or success probability p was calculated by the sample proportion $\hat{p} = x/n$, where n is the number of samples that is obtained from the population. The variable x is the number of observations that possess the particular characteristics of interest that is the number of ECGs in which the model cannot accurately detect the abnormal conditions. With an appropriate n, the sample proportion is assumed to have a normal distribution.

$\hat{p}$ was used to calculate the confidence interval of p, shown as follows:

$$p \in \left(0, 0.05 + 1.645\sqrt{\left(\frac{(0.05)(1-0.05)}{132}\right)}\right) \qquad (1)$$

where 1 – α is a one-sided level of confidence interval for p and z_α is the cumulative distribution function of the standard normal distribution. Therefore, sensitivity can be calculated through the following:

$$\text{Minimum sensitivity rate} = (1 - \text{upper bound of } p) \times 100\%. \qquad (2)$$

VI. Results

Results were determined for both Interoperability and Accuracy (sensitivity, specificity, and overall accuracy). Of the 276 12-lead ECGs used to validate interoperability, the system was able to diagnose each and every ECG with heart rates, heart rhythm, abnormal ECG measurements, and possible diagnostic findings. Therefore, it was concluded that the system amply demonstrated interoperable capability because it was able to present all diagnosis results in a browser from the encoded representation of the open standard HL7 Ontology.

For sensitivity, from 132 ECGs with a variety of abnormal conditions, 125 ECGs were diagnosed correctly according to the physician's diagnosis. The model misdiagnosed 7 ECGs. From (1) and using the critical point $z_\alpha = z_{0.05} = 1.645$, a one-sided 95% confidence interval for the probability p of sensitivity error was as follows:

$$p \in \left(0, 0.05 + 1.645\sqrt{\left(\frac{(0.05)(1-0.05)}{132}\right)}\right)$$
$$p = (0, 0.09)\,.$$

This result shows that the *upper bound* on the probability p is 0.09, which can be used to obtain a lower bound on the complementary 1 – p [18] that being 0.91; thus, the model has a sensitivity rate of *at least* 91%. The proportion of ECGs with abnormal conditions that the model can detect correctly will be *at least* 91%.

For specificity, from 144 normal ECGs, 138 ECGs were correctly diagnosed by the model as normal without finding any abnormal conditions. The model misdiagnosed 6 ECGs. The *upper bound* on the probability p is 0.07 providing a lower bound of 0.93; thus, the model has a specificity rate *at least* 93%. Consequently, the proportion of normal ECGs that the model can interpret correctly will be *at least* 93%.

The overall accuracy of the model was calculated by considering the total number of samples and the total number of errors from sensitivity and specificity tests. From 276 ECGs, the acceptable error rate $(p_2) = 0.10$ does not fall in (0, 0.07). Consequently, the proportion of normal and abnormal ECGs that the model diagnoses correctly will be *at least* 93%.

VII. Discussion

The accuracy of the model is acceptable because the results yielded less than the designated 10% error rate. However, it was seen that the error rate of sensitivity is higher than that of the specificity. The explanation for this result is that it is more difficult to automatically diagnose an ECG that indicates abnormal conditions. If an obvious abnormal condition occurs, the model will diagnose it correctly. However, in some challenging records, when an abnormal condition is embedded within another abnormal condition in the same ECG strip, a misdiagnosis is likely to occur. For example, when an ECG strip has multiple beats of Left Bundle Branch Block, and there is a single beat of Premature Ventricular Contraction occurring between those Left Bundle Branch Block beats, the model will detect Left Bundle Branch Block but may or may not detect Premature Ventricular Contraction.

In some records, physicians' diagnoses were different than the abnormal conditions diagnosed by the model. For example, one ECG record was diagnosed as a normal ECG by a physician. However, the model found "Irregular Rhythm" in this ECG strip. The atrial rate that was measured by P–P duration has slightly varied rates. The model incorrectly diagnosed Irregular Rhythm because it detected a wide range of atrial rates from 54 to 66 beats per minute.

After being tested on the data sets, an exploratory analysis was done by systematically modifying sensitivity and specificity parameters to determine why the model only had 93% total accuracy. This analysis showed a tradeoff. When the model got better at correctly interpreting normal ECGs as normal (higher specificity), it tended to be less sensitive to correctly detecting an abnormal condition when diagnosing abnormal ECGs (lower sensitivity).

VIII. Conclusion

The developed ontology provides a structure for the representation and open exchange of ECG data so that it can be made readily available for viewing on a multitude of computing platforms. In addition, by using ECG digital data as inputs, abnormal cardiac conditions can be automatically diagnosed and presented. Both of these capabilities are based on the HL7 standard for point-of-care medical device communication [7]. The ontology was encoded in XML vocabularies providing human and machine readable formats, allowing it to be displayed in an Internet browser.

This research shows that an open platform and software independent system can be developed to represent, openly exchange, and accurately diagnose cardiac data. ECG data can be shared and diagnosed across systems without the need for proprietary ECG software. It also promotes the use of the HL7 standard for data exchange and provides an automated ECG diagnosis system with acceptable accuracy rates that can be used as a decision aid for ECG diagnosis. The researchers encourage ECG device vendors to provide capability to map output data from their products to an open standard for data exchange. Such technology serves to advance emerging telemedicine and assistive medicine technologies for providing better healthcare to patients.

Future research can improve understanding of ECG diagnosis and develop a more accurate ECG diagnosis model. An example of methods to improve the accuracy rate of the model is to include ECG images as pattern recognition of knowledge for diagnosis. More detailed measurement of the ECG waveform can also be implemented for a wider range of diagnosis. Demographic information such as age and gender can also be included for a more accurate ECG prediction algorithm. Moreover, an inclusion of a preliminary morphological classification with information of beat number and beat type in the current ontology may add benefit.

References

[1] N. Hughes, L. Tarassenko, and S. Roberts, "Markov models for automated ECG interval analysis," in *Advances in Neural Information Processing Systems 16 (NIPS 2003)*, S. Thrun, L. Saul, and B. Schoelkopf, Eds. Cambridge, MA: MIT Press, 2004.

[2] H. Hiroki, K. Arakawa, J. Muramatsu, J. T. Sugimoto, T. Sawayama, K. Inoue, N. Kawai, and T. Mizutani, "New electrocardiographic criteria for diagnosing right ventricular hypertrophy in mitral stenosis-comparison with the Bonner's and Mortara's criteria," *J. Japanese Circulation*, vol. 52, no. 10, pp. 1114–1120, 1988.

[3] B. Heden, O. Hans, R. Rittner, and L. Edenbrandt, "Acute myocardial infarction detected in the 12-lead ECG by artificial neural networks," *Circulation*, vol. 96, no. 6, pp. 1798–1802, 1997.

[4] P. Porela, K. Hanninen, T. Vuorenmaa, M. Arstila, K. Pulkki, A. Bredbacka, K. J. Antila, J. Jalonen, H. Helenius, and L.-M. Voipio-Pulkki, "Computer-assisted electrocardiograhy in structured diagnosis of acute myocardial infarction," *J. Scand. Cardiovasc.*, vol. 33, no. 2, pp. 89–96, 1999.

[5] *Health Informatics. Standard Communication Protocol. Computer-Assisted Electrocardiography,* IST/35 Standard BS EN 1064, 2005.

[6] *Health Level Seven (HL7)*. (2014). [Online]. Available: http://www.hl7.org

[7] *Draft Standard for Health informatics—Point-of-care medical device communication—Nomenclature,* IEEE P1073-1.1.1/d08, pp. 62–81, 2003.

[8] D. Jenkins and S. Gerred (1996). *ECG library.* [Online]. Available: http://www.ecglibrary.com/ecghome.html

[9] F. G. Yanowitz. (2005). *ECG learning center.* [Online]. Available: http://ecg.utah.edu

[10] H. Wang, F. Azuaje, G. Clifford, B. Jung, and N. Black, "Methods and tools for generating and managing ecgML-based information," in *Computers in Cardiology*, Chicago, IL, USA: IEEE, vol. 31, 2004.

[11] Regulated Clinical Research Information Management Technical Committee. *HL7 aECG implementation guide* [Online]. Available: http://en.wikipedia.org/wiki/HL7_aECG

[12] *ECG Interpretation Made Incredibly Easy*, 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins, 2005.

[13] D. Dubin, *Rapid Interpretation of EKG's*, 6th ed. Tampa, FL, USA: Cover Publishing Company, 2000.

[14] H. J. L. Marriott, *Practical Electrocardiography*. Philadelphia, PA, USA: Williams & Wilkins, 1988.

[15] P. Walmsley, *Definitive XML Schema*, 2nd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 2013.

[16] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley. Physiobank, Physiotoolkit, and Physionet: Components of a new research resource for complex physiologic signals. *Circulation* [Online]. 101(23), e215–e220. Available: http://circ.ahajournals.org/cgi/content/full/101/23/e215

[17] ECG Databases. (2014). [Online]. Available: http://physionet.org/physiobank/database/#ecg

[18] A. J. Hayter, *Probability and Statistics for Engineers and Scientists*. Boston, MA: PWS-Kent, 1996.

[19] S. Kaufman, I. Poupyrev, E. Miller, M. Billinghurst, P. Oppenheimer, and S. Weghorst, "New interface metaphors for complex information space visualization: An ECG monitor object prototype," in *Medicine Meets Virtual Reality: Global Healthcare Grid.* San Diego, CA: IOS Press, 1997, pp. 131–140.

[20] U. Ayesta, L. Serrano, and I. Romero, "Complexity measure revisited: A new algorithm for classifying cardiac arrhythmias," in *Proc. 23rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2001, vol. 2, pp. 1589–1591.

[21] H. Burstein, *Attribute Sampling: Tables and Explanations; Tables for Determining Confidence Limits and Sample Size Based on Close Approximations of the Binomial Distribution*. Burr Ridge, IL: McGraw-Hill, 1971.

Thidarat Tinnakornsrisuphap received the B.S. degree in 1998, industrial engineering from Chulalongkorn University, Bangkok, Thailand, and the M.S. and Ph.D. degrees in 2002 and 2006, industrial engineering from Oregon State University, Corvallis, USA.

She graduated from the School of Mechanical, Industrial and Manufacturing Engineering, Oregon State University. Now she resides in San Diego, CA, USA. Her current research interests include information systems engineering, human factors engineering, database management, data modeling, web applications, statistics, and ergonomics.

Richard Billo received the B.A. degree from West Virginia University, in 1978, the M.A. degree from the University of the Pacific, in 1981, both in psychology, and the M.S. and Ph.D. degrees in industrial engineering from Arizona State University, AZ, USA, in 1987 and 1989, respectively.

He is currently an Associate Vice President for Research at the University of Notre Dame, Notre Dame, IN, USA. He is currently a Member of the Editorial Board for the *International Journal of Radio Frequency Identification Technology and Applications*.

He was involved in research for the National Science Foundation, the Department of Defense, the Department of Energy, the U. S. Postal Service, and numerous corporations. His current research interests include wireless networks, bioinformatics, and automatic data capture.