Journal of Pharmacological and Toxicological Methods 81 (2016) 99–106

Contents lists available at ScienceDirect

Journal of Pharmacological and Toxicological Methods

journal homepage: www.elsevier.com/locate/jpharmtox

Unlock the information in your data: Software to find, classify, and report on data patterns and arrhythmias

Anil C. Mehendale *, Jennifer M. Doyle *, Christopher M. Kolin, John P. Kroehle Jr.

Data Sciences International, 119 14th St NW, Suite 100, St. Paul, MN 55112, United States

ARTICLE INFO

Article history:
Received 1 March 2016
Received in revised form 9 May 2016
Accepted 12 May 2016
Available online 16 May 2016

Keywords:
Arrhythmia
Automated analysis
Cardiovascular safety
ECG
Jacketed Telemetry
Methods
Nonclinical
Pattern recognition
Telemetry
Toxicology

ABSTRACT

Introduction: Safety studies generate a significant volume of waveform and calculated data. The verification of calculated data and the process of searching through these data for patterns of interest (including arrhythmias) is time intensive. Data Insights™ has been developed for the Ponemah™ software platform to provide efficient verification and search capabilities.

Methods: Searches may be constructed using calculated and pattern matching data available in Ponemah. Searches are composed of one or more search clauses that may be combined using Boolean operators (AND, OR). Each search clause is a Boolean expression composed of inputs and functions. Data Insights includes a number of predefined species-specific searches for arrhythmias that were qualified for canine, non-human primate and minipigs. Qualification compared arrhythmias identified using Data Insights against a board-certified veterinary cardiologist hand-scored reference datasets.

Results: In seven out of eight arrhythmia types, arrhythmia incidences identified by Data Insights were congruent to those identified by hand-scoring. Premature Atrial Contractions (PACs) accounted for the only discrepancy in hand scored data-segments, although all overt PACs identified by the veterinary cardiologist were also identified by Data Insights. Unscored atrio-ventricular blocks accounted for the remaining differences.

Discussion: Data Insights may be used to support different applications, as searches may be created for any physiologic signal type. Its interactive dialog permits rapid review of search results and a dynamic method for handling outliers, signal noise, and false positives. Data Insights provides an efficient method to locate, present, and report on data patterns and anomalies for accurate, consistent results.

1. Introduction

Over the past few decades, improvements to data acquisition and analysis systems have led to an increase in the volume of physiological data collected for drug development and the level to which these data are being analyzed has increased dramatically. The amount of data for a typical continuous 24-h acquisition in a safety pharmacology study could be on the order of 100,000 cardiac cycles, or more, per animal. The desire to combine safety pharmacology endpoints in toxicology studies increases the number of such acquisitions (Pettit, Berridge, & Sarazan, 2010). Analyzing large datasets adds another burden on the researcher performing the analysis, increasing the analysis time and operational cost while demands for time and financial efficiencies are equally high. Semi-automated analysis methods are available; however, thoroughly auditing analysis results is also time consuming.

The presence of unexpected arrhythmias have been associated with the withdrawal, or restricted use, of many pharmaceuticals (Piccini et al., 2009). Current methods for arrhythmia analysis typically rely on hand-scored or semi-automated analysis on a representative sample consisting of only a few cardiac cycles at selected time point(s) throughout the ECG data and do not provide an efficient or comprehensive detection method.

In response to these needs, Data Sciences International (DSI) developed Data Insights™ for the Ponemah™ software platform. Data Insights allows the researcher to assess the quality of their data analysis and target problem areas for additional cleaning and analysis without having to manually over read the dataset. Data Insights reveals these

Abbreviations: ar_V Ectopic, the "ar_" prefix indicates default Data Insight arrhythmia searches (For example, ar_V ectopic refers to an arrhythmia search for ventricular ectopic cycles); Cyc, cycle is used by Data Insights to identify cycles relative to the current cycle (cyc0); ECG, electrocardiogram; HR, heart rate in beats-per-minute computed by Ponemah as the reciprocal of the RR-I (in seconds) for the cardiac cycle multiplied by 60; JET, Jacketed External Telemetry; LVEDP, left ventricular end diastolic pressure; Num, a sequential number Ponemah assigns to each cardiac/respiratory cycle; PAC, Premature Atrial Contraction; PCt, P wave count representing the number of valid P waves marked by the Ponemah ECG analysis module within a user-defined logging period; PR-I, PR interval is the time interval (in milliseconds) between the beginning of atrial depolarization and the beginning of ventricle depolarization; PVC, Premature Ventricular Complex; QRS width, QRS width is the time interval (in milliseconds) of the QRS complex from the Q wave to the S wave; RR-I, RR interval is the time interval (in milliseconds) from one R wave to the previous R wave; SSD, Solid State Drive; Sys, systolic blood pressure.

* Corresponding authors.

http://dx.doi.org/10.1016/j.vascn.2016.05.007
1056-8719/

problem areas by applying user-defined search rules to the dataset and displaying cycles that match the search criteria. Match results are displayed in graphical and tabular formats to provide researchers a multi-faceted view of each search result and how the results are distributed throughout their dataset. This provides researchers the necessary information to better understand their data and make an informed decision on whether or not to exclude certain sections of data due to signal artifact or data dropout without investing a significant amount of time.

In addition to searches for data validation and analysis, researchers may create Data Insights searches to locate, present, and report on any data pattern or anomaly within the dataset, e.g., cardiac arrhythmias. The use of Data Insights permits the user to move beyond snapshots of data for efficient coverage of large volumes of data.

Using Data Insights searches, researchers can optimize their data review and analysis process to achieve consistent, reproducible results for reporting purposes.

2. Methods

In this section we present the salient features included in Data Insights to permit the design and editing of search definitions and interactive processing of results, Data Insights application to arrhythmia detection, and the steps taken to qualify Data Insights functionality and usability.

2.1. Search definition

Searches are composed of one or more search clauses that may be combined using Boolean operators (AND, OR). Each search clause is a Boolean expression composed of inputs and functions. The inputs include amplitude and timing data calculated by Ponemah (RR interval, heart rate, LVEDP, etc.), pattern matching results, and time of day information. The functions determine how inputs are used when evaluating the Boolean expression that forms a search clause. A few examples are provided in Table 1. Existing search definitions may be modified or new search definitions may be created by the user. One or more search definitions may be associated with each acquired signal. The relationship between raw signals, derived cycle information, and Data Insights is shown in Fig. 1.

2.2. Result processing and threshold determination

The following features assist in the processing of large sets of results and the identification of appropriate thresholds for use in search definitions.

2.2.1. Display of search inputs

Each search input used in a search definition is represented in the Results Derived View as shown in Fig. 2. Each search result is represented by a row in the Results Derived View which displays the values for each search input, providing an understanding of why a search result met the search criteria.

2.2.2. Visualization of results

Each search result is displayed in the Results Wave View separated by a dashed green line. The range of time displayed for each result depends on the range of data that contributes to the result.

Results are also displayed in a histogram as shown in Fig. 3. The width of each histogram bar corresponds to a user defined bin length, which only impacts the visualization of the data. The height of each green histogram bar corresponds to the number of matches in the bin. The red regions indicate the location and number of matches that are currently displayed in the Results Wave View.

If the usable data in a bin drops below the value specified in the Data Percentage field, the background of the bin will change to a coral color (not shown). This alerts the user to regions that have a break in acquired data or have lost a significant amount of data to noise.

2.2.3. Sorting by search input

By default, the search results are ordered by time although the search results may be sorted by any of the columns in the Results Derived View. Sorting updates the order of results in both the Results Derived View and the Results Wave View. In Fig. 2, results are sorted by the %Decrease column.

2.2.4. Bulk handling of results

Multiple results may be selected and either rejected as not belonging to the set of results or marked as bad data to prevent further analysis on the data samples.

2.2.5. Synchronization with graph pages

Each result may be synchronized with standard graph pages and derived listings to provide additional context to the result.

2.3. Arrhythmia detection: Searches

Data Insights includes a number of predefined species-specific searches; canine defaults are listed in Table 2. The default searches are suitable for most subjects. If required, a subject's search may be modified to account for variation specific to the subject. A sample workflow for identifying arrhythmias is also shown in Fig. 4. Clauses related to search result duration and noise sensitivity are not shown.

2.3.1. Ventricular ectopic derivatives

Searches for Premature Ventricular Complexes (PVCs), escape beats, interpolated beats, ventricular bigeminy, trigeminy, couplets, triplets and runs are available as extensions to cycles identified as ventricular ectopics. PVCs, escape beats and interpolated beats are identified by examining the inter-beat intervals preceding and following the ventricular beat. Ventricular couplets, triplets and runs are identified using Series

Table 1
Example Data Insights' search definitions and corresponding cycle identification description illustrating the diversity of searches that can be created using any acquired signal's (ECG, left ventricular pressure, systemic blood pressure, etc.) derived data and ECG waveform morphology using ECG Pattern Recognition Option (PRO).

Search definition	Cycle identification
Value(Sys_{cyc0}) > 170	Identifies all cycles where the systolic value is greater than 170 mmHg.
%Decrease($RR\text{-}I_{cyc\text{-}1}$, $RR\text{-}I_{cyc0}$) > 30	Identifies all cycles that show a decrease in the RR interval of more than 30% from the previous cycle (cyc-1) to the current (cyc0). Functions that are similar to %Decrease() are %Increase(), %Change(), Increase(), Decrease() and Change().
Template(ECG_{cyc0}) = Ventricular Ectopic	Identifies all cycles that match templates tagged with a ventricular ectopic tag.
Search($_{cyc0}$)!= ar_PAC	Identifies all cycles that do not match a PAC search. A Search() function would typically be used in conjunction with other clauses.
RealTime() > 08:00:00 AND RealTime() < 20:00:00	Identifies all cycles that fall between 8:00 AM and 8:00 PM within each 24 h cycle, such a search would typically be used in conjunction with other clauses.
Series(ar_V Ectopic, 1)=2	Identifies multiple occurrences of a pattern. In this example, all instances where exactly two consecutive cycles match an ar_V Ectopic search are identified.