

Modelling Electrocardiogram Waveform Interpretations for SEND Submissions

Jack Baker; Labcorp Early Development Laboratories Limited, Huntingdon, UK; Jack.Baker@labcorp.com

Abstract

Modelling data for electrocardiogram (ECG) waveform interpretations can be challenging in SEND. Currently, within the latest confirmed data endpoints for exchange (CoDEx), counts of arrhythmia types in continuous ECG measurements are not in scope for SEND. This poster explores methods to present counts of arrhythmia types in continuous ECG measurements within the ECG domain and provides best practice recommendations for future SENDIG guidance.

Introduction

Arrhythmia counts within ECG measurements have previously been considered out of scope for SEND Implementation. Electrocardiogram waveform data is collected and reported in a variety of different ways in the nonclinical industry. It can vary based on test site, study type, species or recording methods such as implanted telemetry, jacketed external telemetry or multi-lead recordings. Even studies that are collected using similar methods may then be reported differently. This variance adds challenge to modelling the data into SEND format in an efficient and reproducible way. Despite this, there is still precedent to model this data in SEND format, which will allow for the data to be presented in a standard and uniform way. It will create potential to use the SEND datasets for further cross study analysis and visualizations, which will help us better understand and explore the data.

Modelling in SEND Format

Let’s look at how this data has been modelled in the past. Previously this data was modelled by essentially ‘forcing’ the data into the current domain model. Figure 1 shows an example of how this data can be modelled in the current SENDIG 3.1. EGORRES contains both the count and the arrhythmia type. EGSTRESC contains the arrhythmia type, which is mapped to the ‘ECG Result’ controlled terminology, and EGSTRESN contains the count of arrhythmias. This however is not the most appropriate way to model the data, as it goes against the guidelines for the --STRESN variable, which is supposed to contain a numeric form of --STRESC.

My research led me to discover that the CDISC SEND Safety Pharmacology SubTeam had been working on a model to include arrhythmia counts in SEND. While the model is not yet finalized, a suggested model that includes new SEND variables has been created that allows this data to be presented in SEND format. Figure 2 shows an example of the latest proposed model. In this model, EGORRES and EGSTRESC contain the type of arrhythmia, mapped to the ‘ECG Result’ terminology. EGSTRESN remains blank, and a new variable RLTCNT (Result Count) is introduced, which contains the count of arrhythmias. An EGSEV (Severity) variable is also introduced, which contains the severity of the count mapped to a new controlled terminology codelist. This new codelist will map the result captured in the Result Count variable to a controlled term. For example, a count of 2 would map to the controlled term of “DOUBLE”. The full controlled terminology codelist can be seen in Figure 4.

This proposed model allows the counts to be represented in the domain without ‘forcing’ them into the current domain model. Introducing the new ‘Result Count’ variable into the SENDIG will also allow opportunity for other domain models to potentially introduce a similar variable for other types of data.

EGTESTCD	EGTEST	EGCAT	EGORRES	EGORRESU	EGSTRESC	EGSTRESN	EGSTRESU	RLTCNT	EGSEV
INTP	Interpretation	INTERPRETATION	1 JUNCTIONAL ESCAPE COMPLEXES		JUNCTIONAL ESCAPE COMPLEXES			1	SINGLE
INTP	Interpretation	INTERPRETATION	2 SINUS ARREST/PAUSE		SINUS ARREST/PAUSE			2	DOUBLE
INTP	Interpretation	INTERPRETATION	3 2ND DEGREE AV BLOCK		2ND DEGREE AV BLOCK			3	TRIPLE
INTP	Interpretation	INTERPRETATION	6 PREMATURE VENTRICULAR COMPLEX		PREMATURE VENTRICULAR COMPLEX			6	RUN
INTP	Interpretation	INTERPRETATION	11 PREMATURE ATRIAL COMPLEXES		PREMATURE ATRIAL COMPLEXES			11	RHYTHM

Figure 2. CDISC SEND Safety Pharmacology SubTeam domain example.

Variable Name	Variable Label	Type	Controlled Terms, Codelist, or Format	Role	Notes	Core
RLTCNT	Result Count	Num		Result Qualifier	Count of the number of arrhythmias observed	Perm
EGSEV	Severity	Char	New EGSEV Codelist	Record Qualifier	Severity of the Result Count mapped to controlled terminology	Perm

Figure 3. New proposed variables.

Use Cases and Visualization

My research also led me to investigate the use cases of modelling this data into SEND format. Modelling the data to SEND format allows us to visualize and compare data in a standard way. In particular, I have investigated the use of heat maps to visualize and compare data. Let’s use a simple heat map to compare the overall counts of the Result Count variable for the different occasions collected on a Latin Square study. This is shown in Figure 5. The study contained mock data that was created to show a correlation between dose level and arrhythmia counts; you can see this trend in the heat map as it changes from Green to Red as the dose level rises.

Count of RLTCNT						
	Vehicle Control	Low Dose	Mid Dose	High Dose	Grand Total	
ABC-001	2	10	25	41	78	
ABC-002	1	19	27	56	103	
ABC-003	3	13	23	45	84	
ABC-004	5	15	28	52	100	
Grand Total	11	57	103	194	365	

Figure 5. Heat map for overall arrhythmia count on Latin Square study.

After discussing this internally with colleagues in the Safety Pharmacology team, I discovered that the heat maps for overall count may not always be useful for analysis; instead an overlook of specific arrhythmia types over the 24-hour collection period would be more useful for spotting significant trends in the data. Observing the data subject by subject would also be more useful as each subject may have different background arrhythmias that could affect the analysis. Figure 6 shows individual counts of arrhythmia for a specific subject, split out by arrhythmia type, captured over the 24-hour collection periods. This gives a lot better overview of the whole collection period and it allows the data to be analyzed for each subject’s specific arrhythmia types. It allows us to still see the trend as the dose level rises, but we can also see how the specific timepoints are affected. For example, we can see the arrhythmias are more frequent at around the 3-6 hour timepoint, and then, we see a decline as time goes on. It was also highlighted that the values on the high end of the scale in red for the heat maps may sometimes be misleading. In some cases, if a subject has a high amount of arrhythmias, it may not actually indicate any significant issues or have any impact on the study; therefore, this needs to be considered when analyzing this data. It was also noted that the color scale could be improved upon. The color scheme used in Figures 5 and 6 is based on one study’s data, but for some types of studies, a high arrhythmia count may be more common than other studies; therefore, if the color scheme could be standardized across multiple studies, and then adjusted based on things such as study type, dose route, dose levels, etc., a more useful heat map can be created for cross study comparison.

Conclusions

For studies with ECG data as a focus, modelling this data in SEND format can certainly be useful, and although industry may be some way from collecting this data in a standard way that will allow for more efficient SEND generation, there is value in capturing waveform data in SEND format, which will allow data analysis and visualization.

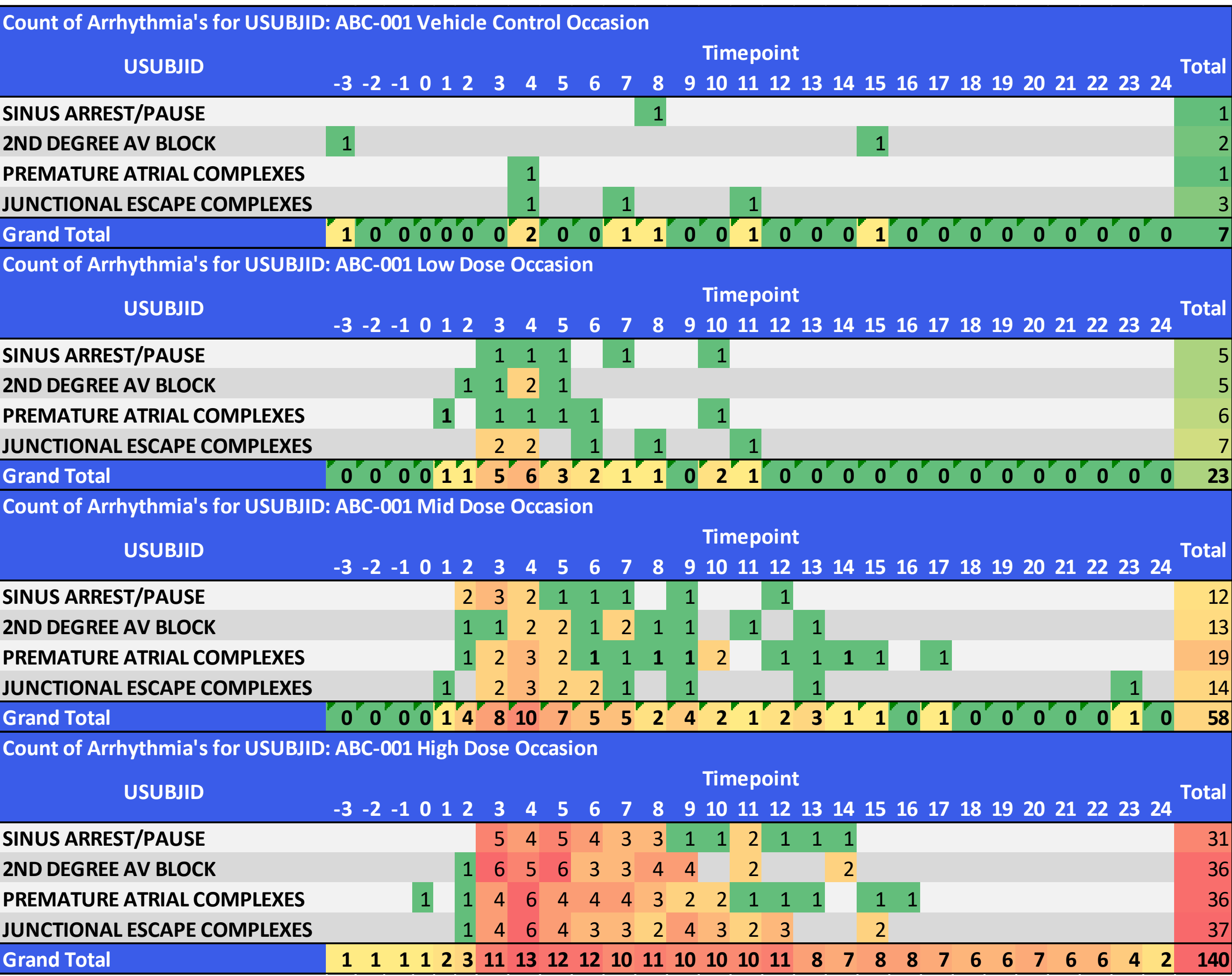


Figure 6. Heat map for Subject ABC-001’s individual arrhythmia types over 24-hour collection period.

Acknowledgements

I would like to acknowledge and thank the Labcorp Safety Pharmacology team who contributed toward the great discussions and exploration of this data, as well as the CDISC SEND Safety Pharmacology SubTeam for their progress and work on the SEND model.