Journal of Pharmacological and Toxicological Methods 92 (2018) 43–51

Contents lists available at ScienceDirect

Journal of Pharmacological and Toxicological Methods

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

Research article

The evaluation of endpoint variability and implications for study statistical power and sample size in conscious instrumented dogs

Alan Y. Chiang[a], Brian D. Guth[b,j], Michael K. Pugsley[c], C. Michael Foley[d], Jennifer M. Doyle[e], Michael J. Engwall[f], John E. Koerner[g], Stanley T. Parish[h,*], R. Dustan Sarazan[i]

[a] Eli Lilly and Company, Indianapolis, IN, United States
[b] Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach, Germany
[c] Purdue Pharma, LP, Stamford, CT, United States
[d] AbbVie, Abbott Park, IL, United States
[e] Data Sciences International, St. Paul, MN, United States
[f] Amgen, Thousand Oaks, CA, United States
[g] FDA, Washington DC, United States
[h] HESI, Washington DC, United States
[i] Independent, Rhinelander, WI, United States
[j] The Preclinical Drug Development Platform, North-West University, Potchefstroom, South Africa

ARTICLE INFO

Keywords:
Double Latin Square design
Myocardial
Left ventricular dP/dt
Variability
Test sensitivity
Heart rate
Time interval average
Super-interval
Statistical power

ABSTRACT

Introduction: The sensitivity of a given test to detect a treatment-induced effect in a variable of interest is intrinsically related to the variability of that variable observed without treatment and the number of observations made in the study (i.e. number of animals). To evaluate test sensitivity to detect drug-induced changes in myocardial contractility using the variable LVdP/dt_{max}, a HESI-supported consortium designed and conducted studies in chronically instrumented, conscious dogs using telemetry. This paper evaluated the inherent variability of the primary endpoint, LVdP/dt_{max}, over time in individual animals as well as the variability between animals for a given laboratory. An approach is described to evaluate test system variability and thereby test sensitivity which may be used to support the selection of the number of animals for a given study, based on the desired test sensitivity.

Methods: A double 4 × 4 Latin square study design where eight animals each received a vehicle control and three dose levels of a test compound was conducted at six independent laboratories. LVdP/dt_{max} was assessed via implanted telemetry systems in Beagle dogs ($N = 8$) using the same protocol and each of the six laboratories conducted between two and four studies. Vehicle data from each study was used to evaluate the between-animal and within-animal variability in different time averaging windows. Simulations were conducted to evaluate statistical power and type I error for LVdP/dt_{max} based on the estimated variability and assumed treatment effects in hourly-interval, bi-hourly interval, or drug-specific super interval.

Results: We observe that the within-animal variability can be reduced by as much as 30% through the use of a larger time averaging window. Laboratory is a significant source of animal-to-animal variability as between-animal variability is laboratory-dependent and is less impacted by the use of different time averaging windows. The statistical power analysis shows that with $N = 8$, the double Latin square design has over 90% power to detect a minimal time profile with a maximum change of up to 15% or approximately 450 mm Hg/s in LVdP/dt_{max}. With $N = 4$, the single Latin square design has over 80% power to detect a minimal time profile with a maximum change of up to 20% or approximately 600 mm Hg/s in LVdP/dt_{max}.

Discussion: We describe a statistical procedure to quantitatively evaluate the acute cardiac effects from studies conducted across six sites and objectively examine the variability and sensitivity that were difficult or impossible to calculate consistently based on previous works. Although this report focuses on the evaluation on LVdP/dt_{max}, this approach is appropriate for other variables such as heart rate, arterial blood pressure, or variables derived from the ECG.

* Corresponding author at: 1156 15th St., N.W., Suite 200, WA 20005, United States.
E-mail address: sparish@hesiglobal.org (S.T. Parish).

https://doi.org/10.1016/j.vascn.2018.02.009
Received 20 December 2017; Received in revised form 30 January 2018; Accepted 28 February 2018
Available online 02 March 2018

1. Introduction

Safety pharmacology studies are conducted on drug candidates to assess for safety relevant effects when administered at therapeutically relevant or higher doses (ICH S7A, 2001). The assessment of possible effects on the cardiovascular system are frequently conducted in conscious dogs that have been chronically instrumented for the collection of the cardiovascular variables of interest using telemetry which typically includes arterial blood pressure, left ventricular pressure and the electrocardiogram (ECG). The maximal rate of pressure increase in the left ventricular during systole ($LVdP/dt_{max}$) has been shown to be a sensitive variable to assess drug-induced effects on cardiac contractility (Guth et al., 2015). Drugs with both positive (amrinone and pimobendan) and negative (atenolol and itraconazole) inotropic effects, known to produce such effects clinically, were tested in a cross-laboratory evaluation and $LVdP/dt_{max}$ proved to be a robust variable to detect dose-dependent effects of the agents tested. For those studies, each of the laboratories included 8 dogs and studies were conducted using a double Latin square design. The use of 8 dogs was based on the extensive experience of the investigators and limited published data with this type of model; however, ultimately the number of animals for the Health and Environmental Sciences Institute (HESI) supported study was selected subjectively.

With each of the four test compounds studied, one treatment arm was the vehicle used without test article. This is an important treatment arm since the vehicle treatment data was used in this study to evaluate the variability of the collected data within and between animals and across laboratories. We propose herein a methodology for making this assessment that should allow any laboratory to determine the variability of all measured variables. Here we report the evaluation on $LVdP/dt_{max}$, but this approach is appropriate for other variables such as heart rate (HR), arterial blood pressure (BP), or variables derived from the ECG. By defining the variability of each variable assessed, the experimenter can define the test sensitivity of their experimental setting in order to answer the question: what size of a drug-induced effect could have been detected? This is of particular importance for studies concluding that no drug-induced effect was found. Furthermore, since the test sensitivity is also a function of the number of animals included in a study, this approach provides a rational approach for deciding how many animals to include in such a study. This is often mandatory for research scientists to obtain permission from either Institutional Animal Care and Use Committee (IACUC) or governmental agencies (such as the National Institutes of Health, NIH) to conduct this type of non-clinical study.

2. Materials and methods

2.1. Test facilities

Studies were performed by 6 independent companies and data were reported previously (Guth et al., 2015). Each individual study was subject to the local guidelines in terms of the vivarium conditions, study conduct and animal use approval procedures. All participating institutions have warranted strict adherence to all applicable animal use regulations in the conduct of these studies. Although efforts were made to harmonize testing procedures and conditions, the local animal use regulations were always prioritized should any conflicts have arisen during the conduct of the study.

2.2. Experimental animals

All participating laboratories used purpose bred beagle dogs acquired from a vendor within their geographic region (North America or Europe). Some laboratories used only male dogs and other laboratories used both males and females. The source and sex of the dogs used by the various laboratories were reported previously (Guth et al., 2015).

Most animals had been used previously during the conduct of safety pharmacology studies but were healthy and free of any residual test article at the start of the study. At one laboratory the animals were naïve at the study onset. No animals were required to be euthanized in the context of this study. After an appropriate recovery period following surgery or washout period after receiving a drug, animals were subjected to a standard clinical pathology examination to evaluate their health status according to local procedures (typically including blood cell counts, serum electrolytes and biochemistry parameters indicative of kidney and liver function) and were qualified for use in further studies.

2.3. Telemetry instrumentation

Each participating laboratory used one of three commercially available implantable large animal telemetry systems; PhysioTel™ model D70-PCTP (Data Sciences International, St. Paul, MN), PhysioTel™ Digital model L21 (Data Sciences International, St. Paul, MN), or ITS model T27 (Konigsberg Instruments, Monrovia, CA).

Regardless of the telemetry system used, all dogs were instrumented to monitor aortic BP, left ventricle pressure (LVP), the ECG, body temperature and activity. Note, however, that body temperature and activity endpoints were not evaluated during the conduct of the study. All methods related to the surgical preparation of animals, telemetry implants and recording systems employed, and drugs evaluated are found in Guth et al. (2015) and Pugsley et al. (2017).

2.4. Study design

Four different treatments were administered to each dog in the order prescribed by a randomly generated double Latin square design over four treatment days at each test site with an appropriate washout period between days (Guth et al., 2015). The washout period was a minimum of 72 h between treatment days. The double Latin square study design combines two randomly generated 4×4 Latin squares (Sarazan et al., 2011). See Appendix A for an illustration of Latin square designs.

The food provided was withdrawn approximately 2 h before dosing in the morning and reintroduced in the afternoon, which was well after the anticipated time to peak drug concentration (T_{max}) of the tested drug. The study dosing technicians were not blinded to treatment; however, the studies were conducted by the same technicians within each laboratory under standard GLP procedures. Best practices for animal handling were implemented to minimize any potential bias in telemetry data collection and analysis.

2.5. Data collection and analysis

2.5.1. Raw data (signals)

Digital LVP, aortic BP and ECG signals were continuously acquired from at least one hour prior to dosing through 24 h post dose on each study day. Sampling rates were ≥ 500 Hz for LVP and ECG signals and ≥ 250 Hz for BP signals which is adequate for the frequency content of each of these signal types (Sarazan, 2014). Digital raw data files were archived to electronic media and retained at each individual study site for future analysis as agreed upon within the HESI Cardiac Safety Technical Committee.

2.5.2. Derived data (variables)

Various derived variables were calculated from output of digital acquisition units at each study site. However, for the purpose of this evaluation, only $LVdP/dt_{max}$ data were used. A similar evaluation could be performed with any of the additional variables measured as previously reported (Pugsley et al., 2017).

Derived data were calculated for every cardiac cycle and the results were collapsed into 10-min mean values for analysis. These mean