Safety Evaluation Method of QT Interval and Proarrhythmic Potential in Drug Development

Tadashi Koga

Thesis for the degree of doctor of philosophy
Presented to the Department of Statistical Science,
School of Multidisciplinary Science,
The Graduate University for Advanced Studies, Japan
2015

TABLE OF CONTENTS

1. Introduction

1.1. Background

Torsade de Pointes (TdP) is a specific, potentially fatal arrhythmia associated with prolongation of the QT interval on the surface electrocardiogram (ECG), whereas the QT interval represents the duration of ventricular depolarization and subsequent repolarization. TdP is a form of irregular ventricular tachycardia that is often self-limiting with a varied clinical presentation. The arrhythmia may present with no symptoms, or with palpitations, dizziness, syncope, or sudden death.

From early 1990s to early 2000s, a number of non-cardiovascular drugs were withdrawn from the market in the US or UK, due to reports of sudden cardiac death, arrhythmias and QT interval prolongation (Yap and Camm, 2003). Some drugs withdrawn from markets in the US/UK are shown in Table 1.

Terfenadine (seldane) was used to treat allergies, hives (urticaria), and other allergic inflammatory conditions. For terfenadine, 21 deaths were reported in the UK and 51% percent of these were cardiac related. The drug was withdrawn from the UK market in 1998. Seldane were also removed from the US market and the Canadian market by their manufacturer in late 1997 and 1999, respectively.

Table 1: Drugs Market Withdrawals in the US/UK

Drug	Therapeutic Class	Year of Withdrawal
Prenylamine	Antianginal	1988
Terodiline	Urinary incontinence	1991
Sparfloxacin	Antibiotic	1996
Terfenadine	Antihistamine	1998
Sertindole	Antipsychotic	1998
Astemizole	Antihistamine	1999
Grepafloxacin	Antibiotic	1999
Cisapride	Gastro-esophageal reflux	2000
Droperidol	Schizophrenia	2001
Levacetylmethadol	Opiate addiction	2003

Terfenadine is no longer available for prescription in the UK. Sertindole was an antipsychotic medication, and it was approved and marketed in 19 countries in Europe from 1996. The manufacturer applied for US Food and Drug Administration (FDA) approval in 1996, but withdrew this application in 1998 following concerns over the increased risk of sudden death from QTc prolongation. In a trial of about 2000 patients on taking sertindole in the US, 27 deaths, including 13 sudden deaths, were reported and 58% of these were cardiac related. In Europe, the marketing authorization was suspended by the European Medicines Agency in 1998 and the drug was also withdrawn from the market voluntary.

QT prolongation has been also reported for oder tricyclic antidepressants

(amitriptylin/elavil, imipramine), flourquinoline antibiotics (moxifloxacin and grepafloxacin) and some opioids (methadone, LAAM).

Congenital syndromes of syncope and sudden death were described before the 1990s. Subsequently, several types of congenital long QT syndrome were identified. Researches into the genetics of these syndromes revealed that each syndrome affected a different ion channel. Drug induced QT prolongation and TdP are always associated with inhibition of the human ether related a go-go gene (hERG) or lkr ion channel.

From middle of 1990s, regulatory agencies realized new drug safety testing guidelines were needed to protect patients from cardiac arrhythmias. European Agency for the Evaluation of Medicinal Products (EMEA), released "Points to consider document on non-cardiovascular medicinal products" in 1997. And, European Society of Cardiology holds conference: The Potential for QT Prolongation & Proarrhythmia by Non-antiarrhythmic Drugs in 1999. Health Canada submitted a concept paper (Health Canada, 2001) to FDA in the US, and finally, the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) started working on guidelines in 2001. In 2005, in the US and EU, ICH issued two guidelines; ICH S7B: Nonclinical Evaluation of Potential for Delayed Ventricular Repolarization (QT Interval Prolongation) by Human Pharmaceuticals (ICH

Expert Working Group, 2005), and E14: Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs (ICH Expert Working Group, 2005).

The resultant regulatory response to the withdrawal of these drugs has taken over a decade to mature and was finally codified by the International Committee on Harmonization (ICH).

1.2. Electrocardiogram (ECG)

Torsade de Pointes (TdP) is considered to relate to prolongation of the QT interval. The QT interval represents the duration of ventricular depolarization and subsequent repolarization. In cardiology, the QT interval is a measure of the time between the start of the Q wave and the end of the T wave in the heart's electrical cycle and often measured using an electrocardiogram.

An electrocardiogram (ECG) is a graphical recording of the electrical activity of the heart over a period of time using electrodes placed on a human's or animal’s body. In humans, it is recorded at the body surface and illustrated as a series of waveforms that show voltage over time. The standard 12-lead ECG is useful in humans as an independent marker of cardiac disease and reflects abnormalities of the heart that can arise from

anatomic, hemodynamic, molecular, ionic, and/or drug-induced causes.

The heart is a specialized muscle consisting of four chambers: the right and left atria, and the right and left ventricles. The right atrium receives blood from the body, the right ventricle discharges blood to the lungs, the left atrium receives blood from the lungs, and the left ventricle discharges blood to the body.

The sequence of cardiac activity consists of a period of relaxation called diastole, during which the heart fills with blood, followed by a period of contraction called systole. Alternating contraction and relaxation of the muscle produces the mechanical pumping action of the heart that carries blood around the circulatory system.

The contraction of the heart is driven by small amounts of electrical activity in the heart muscle cells. Certain muscle cells are organized into a specialized conduction system, much as a set of wires would carry electrical impulses. This electrical activity must be accurately timed and coordinated in the various parts of the heart muscle in order for contraction and relaxation to occur properly.

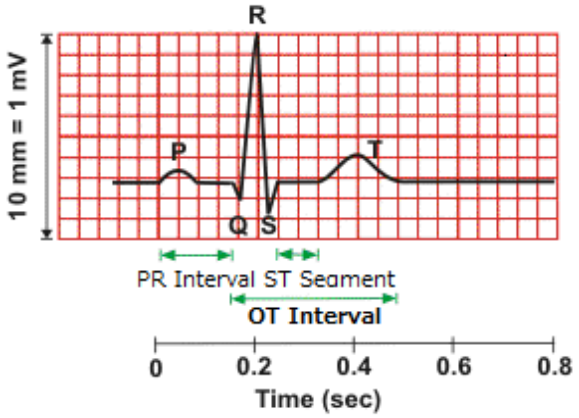

Figure 1: ECG Waveforms

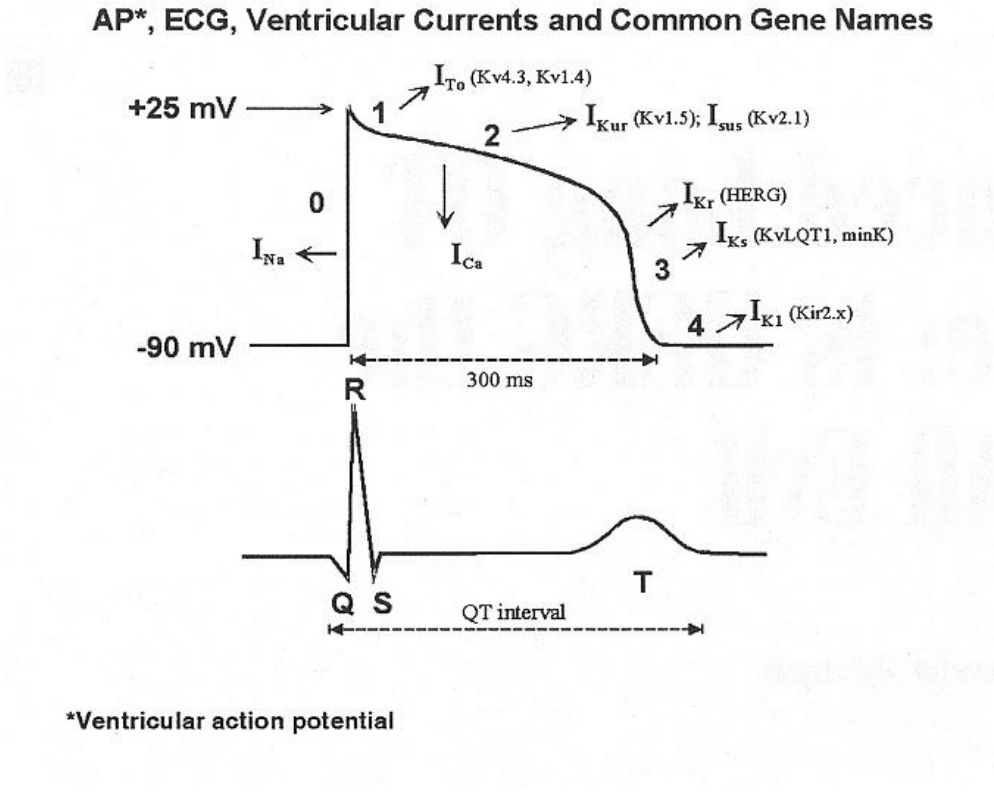

Figure 2: Ion channels involved in cardiac action potentials (Wilt, *et al.*, 1993)

As shown in Figure 1, electrical activities in differing components of the specialized conduction system are responsible for different components of the surface ECG waveforms. The multiple action potentials give rise to the surface ECG waveform. An ECG waveform has single electrophysiological events and combinations of electrophysiological events.

The ECG waveform complex and the measurement points for various intervals are also shown in Figures 1 and 2. There are no ECG components representing atrial repolarization since this occurs at the same time as ventricular depolarization. Well-known ECG waveform components are shown in Table 2.

Usually, heart rate (HR) represents the number of times the heart beats per minute. HR usually implies ventricular rate in the ECG. That means HR implies the number of QRS complexes in a minute. HR also refers to atrial rate, which means the number of P-waves per minute. The common method used ECG machine algorithm derives the median elapsed time from the peaks of the R-waves on the dominant beat type within the 10-second recording. Then HR can be derived as HR = 60,000/RR using RR interval which is measured in msec.

Table 2: ECG waveform components

Waveform component	Description
PR interval	Time from onset of atrial activation to onset of ventricular activation. The PR interval actually ends at the beginning of the QRS complex.
QT interval	Duration of ventricular activation and recovery; includes the QRS, ST segment, and T-wave.
QRS complex	Ventricular depolarization; obscures atrial repolarization. Q-wave is the first negative wave, R-wave is the first positive wave, and S-wave is the first negative wave after R-wave in QRS complex.
ST segment	Delay between ventricular depolarization and repolarization; all ventricular cells are depolarized.
T-wave	Ventricular repolarization.
U-wave	This is a wave on an electrocardiogram that is not always seen. It is typically small, and, by definition, follows the T wave. U waves are thought to represent repolarization of the papillary muscles or Purkinje fibers.
RR interval	Ventricular rate; time between cycles of depolarization/ repolarization of the entire heart.

The PR interval and QRS complex are often of interest because these components imply muscle depolarization. The QT interval is mainly determined by the ventricular repolarization time, including the delay after depolarization. Ventricular depolarization usually found a small component of the QT interval. However, ventricular repolarization time is very important. If this time is long, ventricular depolarization might begin to occur

before repolarization is complete. This will cause the electrical activity of the heart to become unstable, and it might result in loss of effective cardiac function and death.

This abnormal electrical activity is referred to as ventricular tachycardia. One of ventricular tachycardia is known as Torsades de Pointes (TdP). Regrettably, risk factors of ventricular tachycardia including TdP and death are not clear and there are still many unanswered questions about assessing the risk of tachycardia. A pattern of prolonged and shortened QT intervals typically occurs before the ECG change toward Torsade de Pointes. Therefore, the QT interval has been identified as a safety biomarker for the assessment of the risk of TdP. However, prolongation of the QT interval is not always proarrhythmic; the clinical conditions for proarrhythmia are described in Yap and Camm, 2003.

1.3. QT/QTc Interval

The QT interval is considered identified as a safety biomarker for the assessment of the risk of ventricular tachycardia including TdP and sudden death. Ventricular depolarization usually found a small component of the QT interval. However, ventricular repolarization time will extensively shorten as heart rate increases, and extensively lengthen as heart rate decreases.

Currently, it is considered important to correct the QT interval for the HR or RR interval to compare QT intervals that were measured at different HR/RR intervals. The QT interval corrected for HR/RR interval is called as the corrected QT (QTc) interval.

QTc interval is considered to be useful when we compare QT intervals collected at different HR/RR intervals within an individual and populations. Currently, QTc is used as a standard endpoint to assess the potential for a drug to effect cardiac repolarization, although QTc is recognized as an imperfect endpoint to evaluate risk of fatal arrhythmia because there are many risk factors to fatal arrhythmia (Lipicky, 1993; Dmitrienko *et al.*, 2005). Despite the imperfection as an endpoint, QTc is still a common endpoint because there is a qualitative relationship between QT prolongation and the risk of ventricular tachycardia, especially for drugs that cause substantial prolongation of the QT interval.

Methods to correct QT interval from RR interval change will be discussed in later section.

1.4. Safety Pharmacology Study

Typically, non-clinical pharmacology studies can be divided into three type of studies: primary pharmacodynamics, secondary pharmacodynamics, and safety pharmacology studies. Primary pharmacodynamics effects will be studied on the model

of action or effects of a substance in relation to its desired therapeutic target. Secondary pharmacodynamics effects will be studies on the mode of action or effects of a substance not related to its desired therapeutic target. Safety pharmacology studies are studies to investigate the potential undesirable pharmacodynamics effects of a substance on physiological functions relating to exposure of experimental drug in the therapeutic range or above.

In 2005, ICH released ICH S7B guideline; "The Non-Clinical Evaluation of the Potential for Delayed Ventricular Repolarization (QT Interval Prolongation) by Human Pharmaceuticals" with ICH E-14 guideline. As in the ICH S7B guideline, it describes a nonclinical testing strategy for assessing the potential of a test substance to delay ventricular repolarization. (Figure 3). This guideline also includes information concerning non-clinical assays and an integrated risk assessment. This guideline extends and complements the "ICH Guideline on Safety Pharmacology Studies for Human Pharmaceuticals" (ICH S7A) (ICH Expert Working Group, 2000). This guideline also applies to new chemical entities for human use and marketed pharmaceuticals, when appropriate. Pharmaceuticals for which testing is not called for are described in ICH S7A.

Principles and recommendations described in ICH S7A also apply to the studies conducted in accordance with the ICH S7B guideline. The investigational approach and

evidence of risk should be individualized for the test substance, depending on its pharmacodynamic, pharmacokinetic and safety profiles. Generally, in vitro and in vivo assays are complementary approaches; therefore, currently, both assay types would be conducted, prior to the first clinical trial in humans.

Most pharmaceuticals companies learn and confirm to investigate QT prolongation for new drug candidates stating with non-clinical evaluations (Shah, 2002; Garnett *et al.*, 2012).

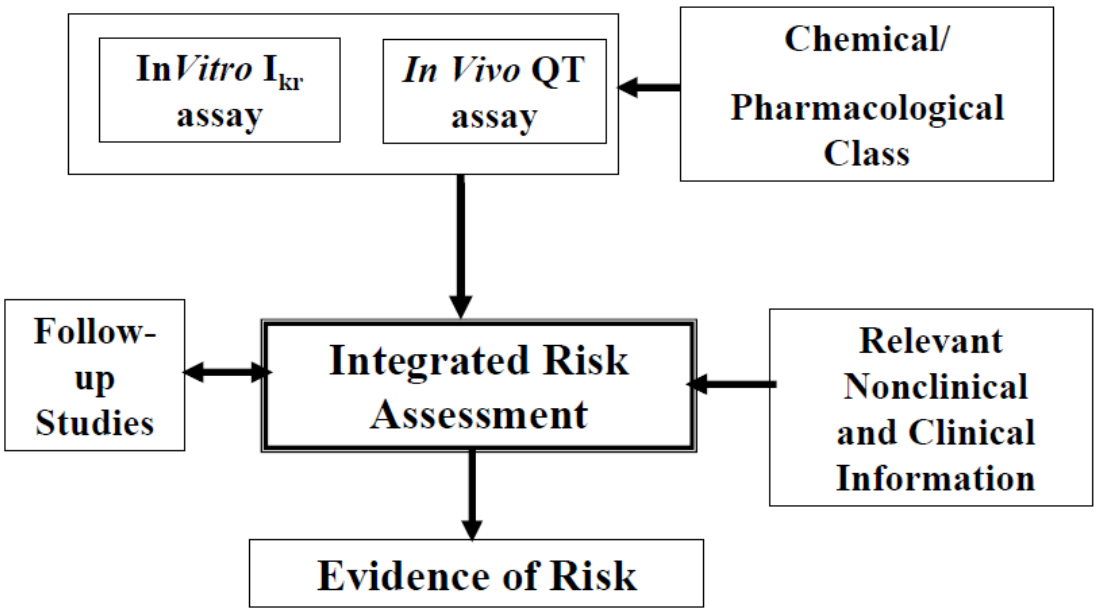

Figure 3: Non-Clinical Testing Strategy (ICH S7B, 2005)

1.5. Cynomolgus Monkeys

Cynomolgus monkeys, one of a number of primates, are frequently used in non-clinical studies due to their close evolutionary and phylogenetically relationships to humans (Fukase *et al*., 1994; Ehara *et al*., 1993). Developmental and reproductive toxicology testing in cynomolgus monkeys has been studies due to the increasing number of biopharmaceuticals in drug development, since this animal is the only species to express pharmacologic responses similar to humans (Chellman *et al*., 2009; Ema *et al*., 2009; Kashima *et al*., 2008). Furthermore, genes in cynomolgus monkeys have been investigated for the species difference occasionally seen in drug metabolism between monkeys and humans (Uno *et al*., 2009; Uno *et al*., 2010).

Cardiac function in cynomolgus monkeys has been well evaluated real time three-dimensional echocardiography (Tsusaki *et al*., 2009), and the procedure to obtain ECGs have been developed for measuring parameters required according to the ICH S7B guideline (Hayes *et al*., 1994; David *et al*., 2006).

2. Reference Intervals in Cynomolgus Monkeys

In medicine, biology, and other healthcare related subjects, a reference interval is the range of values for a physiologic measurement in healthy subjects. For a physician, researcher, or other healthcare professionals, it is often used for comparison to interpret a set of test results for a particular patient or subject. In a target population, a reference interval generally denotes the one in healthy individuals or subjects. Even in drug development process, reference intervals are very popular decision support tool used for interpretation of drug safety results. Here, I consider this fundamental safety evaluation tool in cynomolgus monkeys.

2.1. Cynomolgus Monkeys Background Data

Reference intervals, sometimes called reference ranges or normal ranges, are important in clinical and non-clinical studies to evaluate health condition indices and improve the accuracy of studies to evaluate drug safety. Extensive information on drug safety evaluation in humans is available due to the large number of studies performed (NCCLS, 2000); however, no information is available on reference intervals in hematological and serum biochemical parameters in cynomolgus monkeys (Macaca fascicularis), one of a number of primates phylogenetically close to humans. The

cynomolgus monkey has been adopted as a study animal more frequently in recent years to evaluate drug safety pharmacology, therefore fundamental information on hematological, serum biochemical, and ECG parameters in cynomolgus monkeys should be determined as soon as possible. It has been reported that, with humans, these variables vary with physiological condition (Demacker *et al*., 1982; Evans and Laker, 1995; Widjaja *et al*., 1999), and it is necessary to verify whether similar variation occurs in cynomolgus monkeys.

We conducted a non-clinical study using cynomolgus monkeys to obtain information of this animal. Using data obtained from this study, as background data, summary statistics and reference intervals of hematological, serum biochemical, and ECG parameters in cynomolgus monkeys were calculated to grasp distributions of these parameters. Next, within-animal variation and its ratio to total variation were estimated for each parameter to assess physiological variations in individual animals. Finally, a simple method that applies prior information was proposed to estimate individual reference intervals. This method is useful for evaluating individual health condition index and drug safety in individual animals at evaluation, on the basis of within-animal variation.

2.2. Study System

2.2.1. Animals

Ninety-five male and 95 female Chinese-bred cynomolgus monkeys aged 3 to 7 years were allocated. The animals were acclimated for approximately 1 week. During this period, clinical signs were observed at least once daily and hematological and serum biochemical parameters were measured once. No animal showed abnormalities during the acclimation period, thus data from these animals obtained from 2003 to 2004 were used. Some parameters were not measured in certain animals.

2.2.2. Animal Maintenance Conditions

Room controls were set to maintain temperature and humidity in the ranges of 23˚C to 29˚C and 35% to 75%, respectively, with air changes 15 times/h and a 12-h artificial light cycle. The animals were provided approximately 108g (12g ×9 biscuits) of solid food at approximately the same time daily, and any remaining food was removed on the following day. Drinking water, certified to meet the Water Quality Standard required by the Japanese Waterworks Law, was available ad libitum from an automatic supply.

2.2.3. Hematological and Biochemical Parameters

Parameters, abbreviations, units, and analysis methods are shown in Tables 3 and 4. For hematological analysis, whole blood was drawn from the femoral vein with an anticoagulant. Coagulation tests were performed with 3.8% sodium citrate-treated plasma.

PT and APTT were measured using coagulation analyzer CA-5000 (Sysmex, Kobe, Japan) (Kurata and Horii, 2004). Other hematology parameters were determined or counted with an automated blood cell counter ADVIA 120 (Bayer Medical, Tokyo, Japan) (Moritz *et al*., 2004). For serum biochemical analysis, blood was drawn from the femoral vein, and serum was obtained by centrifugation after stabilization at room temperature for 20 to 60 min. Blood biochemical parameters were measured using automated analyzer BM-8 (JEOL, Tokyo, Japan). All hematological and blood biochemical analyses were conducted four times a month.

Histograms of male cynomolgus monkeys are shown in Figures 4 and 5. Histograms and nonparametric density lines show that many parameters including WBC, leukocyte fractions (Ret, Eosino, Baso, Mono, Lymph, LUC), AST, ALT, ALP, LDH, CPK, T.Bil., TGL, and GLU had a right-heavy tail distribution, and it appears that it would be difficult to assume that these parameters follow normal (gaussian) distribution. The histograms and nonparametric density lines of other parameters, except for ALP, looks to have a

unimodal distribution and could be assumed to have a normal distribution approximately.

Histograms for female cynomolgus monkeys are not shown, but their readings were similar to those of males. Each parameter in males and females was considered to follow a similar distribution for each parameter.

Table 3: Parameters and analytical methods (Hematology)

Abbreviation	Parameter	Unit	Method
RBC	Number of red blood cells	1012/L	Dual angle laser flow-cytometric measurement
WBC	Number of white blood cells	109/L	Dual angle laser flow-cytometric measurement
Ht	Hematocrit value	%	Calculation
Hb	Hemoglobin concentration	g/dl	Modified cyanmethemoglobin method
Platelet	Number of blood platelets	109/L	Dual angle laser flow-cytometric measurement
MCV	Mean corpuscular volume	fL	Dual angle laser flow-cytometric measurement
MCH	Mean corpuscular hemoglobin	pg	Calculation
MCHC	Mean corpuscular hemoglobin concentration	g/dl	Calculation
Ret	Reticulocyte ratio	%	Laser flow-cytometric measurement with RNA stain
Eosino	Eosinophilic leukocyte ratio	%	Flow-cytometric measurement with peroxidase stain an dual-angle laser
Baso	Basophilic leukocyte ratio	%	Flow-cytometric measurement with peroxidase stain an dual-angle laser
Mono	Monocyte ratio	%	Flow-cytometric measurement
Lymph	Lymphocyte ratio	%	Flow-cytometric measurement
Neutro	Neutrophilic leukocyte ratio	%	Flow-cytometric measurement
LUC	Large unstained cell ratio	%	Flow-cytometric measurement
PT	Prothrombin time	s	Quick method
APTT	Activated partial thromboplastin time	s	Quick method

Table 4: Parameters and analytical methods (Blood chemistry)

Abbreviation	Parameter	Unit	Method
AST	Aspartate aminotransferase	IU/L	UV rate assay
ALT	Alanine aminotransferase	IU/L	UV rate assay
ALP	Alkaline phosphatase	IU/L	UV rate assay
LDH	Lactate dehydrogenase	IU/L	WrÓblewski-La Due method
CPK	Creatine phosphokinase	IU/L	UV rate assay
T.Bil.	Total bilirubin	mg/dl	Vanadate oxidation method
TP	Total protein	g/dl	Biuret method
ALB	Albumin	g/dl	Bromo cresol green method
T.Cho.	Total cholesterol	mg/dl	COD·HDAOS[a]-method
TGL	Triglyceride	mg/dl	GPO·HDAOS[b]-method, glycerol blanking
GLU	Glucose	mg/dl	Hexokinase·G-6-PDH[c]-method
BUN	Blood urea nitrogen	mg/dl	Urease-GlDH[d]-method
CRNN	Creatinine	mg/dl	Creatininase·F-DAOS[e]-method
IP	Inorganic phosphorus	mg/dl	PNP·XDH[f] -method
Ca	Calcium	mg/dl	MXB[g]-method

a) COD·HDAOS: Cholesterol oxidase·N-(2-hydoxy-3-sulfopropyl) 3,5-dimethoxyaniline.

b) GPO·HDAOS: Glycerol-3-phosphate oxidase ·N-(2-hydoxy-3-sulfopropyl) 3,5-dimethoxyaniline.

c) G-6-PDH: Glucose-6-phosphate dehydrogenase.

d) GlDH: Glutamate dehydrogenase.

e) F-DAOS: N-ethyl-N-(2-hydroxy-3-sulfopropyl-3,5-dimethoxy-4-fluoroaniline.

f) PNP· XDH: Purine nucleoside phosphorylase · xanthine dehydrogenase.

g) MXB: Methyl xylenol blue.

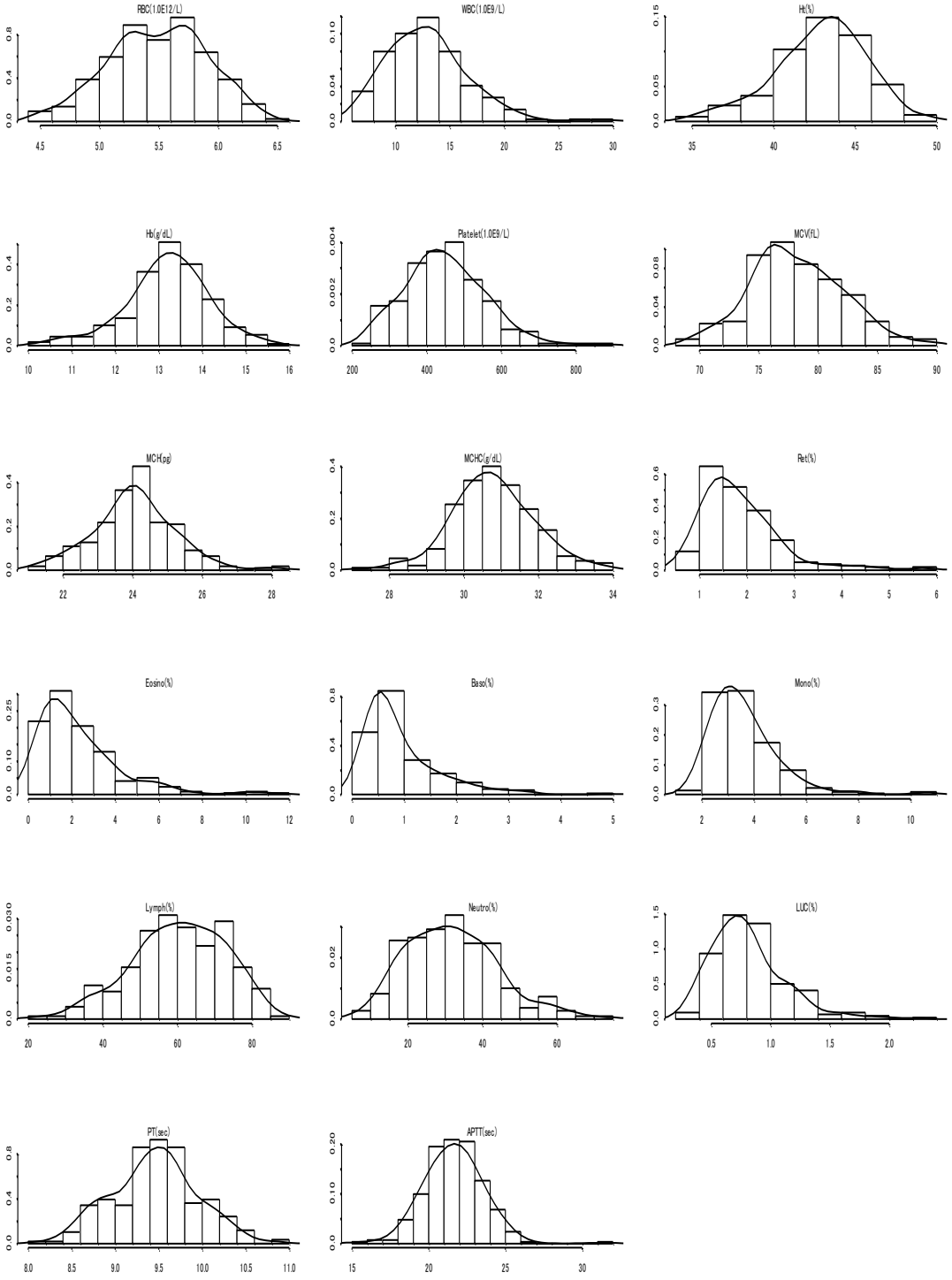

Figure 4: Hematology in male cynomolgus monkeys

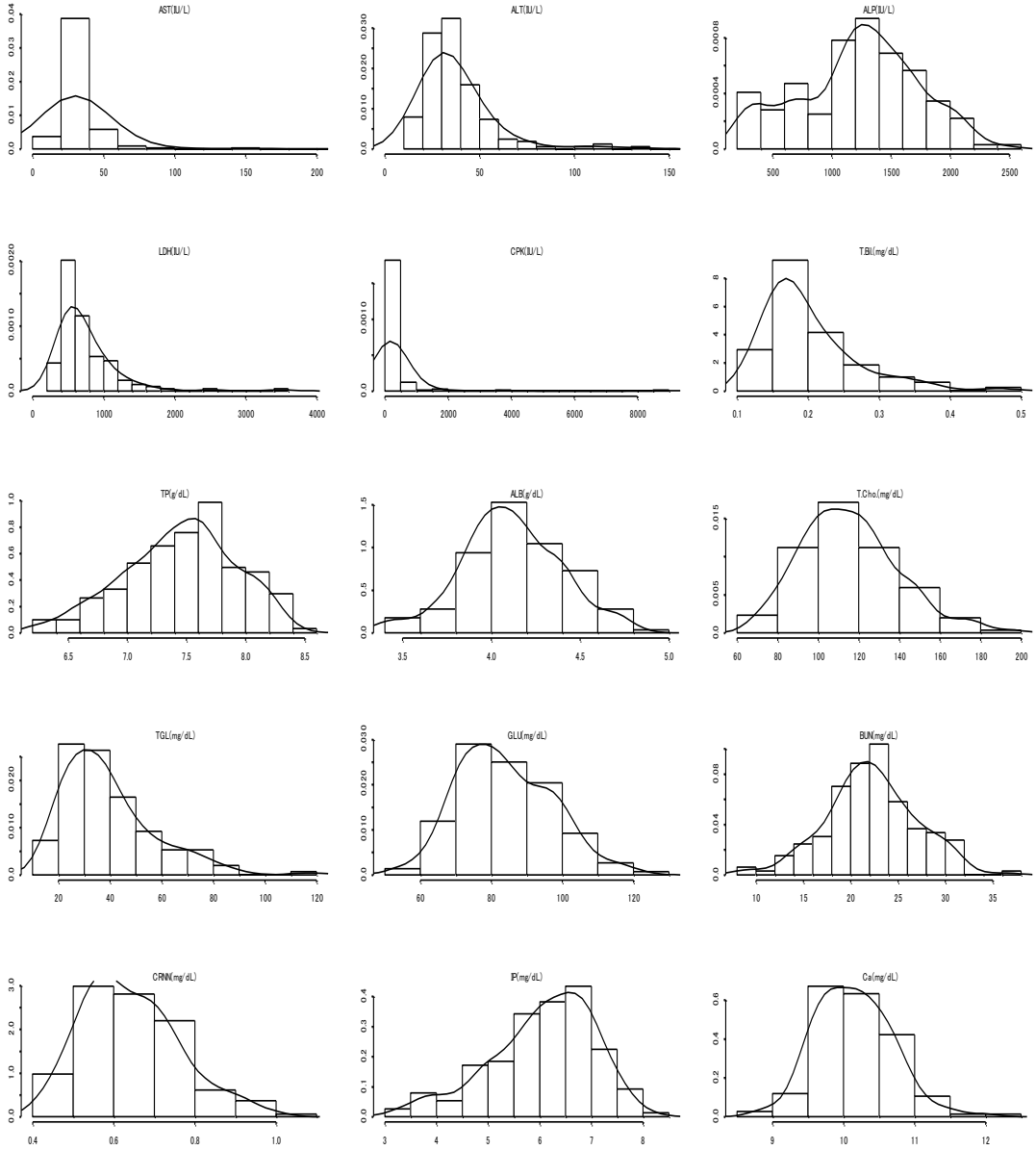

Figure 5: Blood chemistry in male cynomolgus monkeys

2.3. Reference Intervals

Because data were obtained from a relatively large number of animals, classical but simple methods were used for estimating reference intervals (Kaputo, 1972; Amador, 1974). When the data were considered to have followed a normal distribution based on a density plot and histogram, the interval was calculated as mean ± 1.96 SD. When the data were considered to have followed a log-normal distribution, the interval was calculated as mean ± 1.96 SD for log-transformed data and the estimates were then back-transformed to the original scale. Nonparametric percentile estimates (2.5%, 97.5%) (Massod, 1977) were also calculated for comparison. The results for males and females are shown separately in Tables 5 and 6.

All reference intervals and nonparametric reference intervals were considered to be substantially consistent. From these statistics and histograms, a normal/log-normal distribution could be assumed, to which many common statistical analysis methods can be applied.

Table 5: Reference intervals in male cynomolgus monkeys

				Reference Interval		Percentile		Estimates	
Parameter	**N**	**Mean**	**SD**	**UCL**	**LCL**	**97.5%**	**2.5%**	σ^2	β
Hematology									
RBC	220	5.492	0.413	6.307	4.678	6.235	4.645	0.0649	0.6197
WBC*	220	12.916	3.724	21.685	7.048	20.855	6.785	0.0214	0.7189
Ht	220	42.88	2.82	48.44	37.32	47.35	36.25	3.914	0.5083
Hb	220	13.20	0.95	15.07	11.32	15.05	10.85	0.375	0.5870
Platelet	220	450.9	109.0	665.7	236.2	674.5	259.0	2319.68	0.8046
MCV	220	78.21	3.92	85.94	70.48	86.20	70.65	2.344	0.8477
MCH	220	24.05	1.19	26.39	21.72	26.35	21.60	0.188	0.8666
MCHC	220	30.77	1.10	32.93	28.61	33.00	28.35	0.678	0.4355
Ret*	204	1.85	0.85	3.92	0.77	4.29	0.80	0.132	0.3647
Eosino*	220	2.32	1.92	8.32	0.32	7.05	0.30	0.176	0.7181
Baso*	220	0.9	0.7	3.1	0.2	2.8	0.2	0.57	0.2153
Mono*	220	3.6	1.3	7.0	2.0	6.7	2.0	0.13	0.5165
Lymph*	220	60.40	12.56	85.68	35.96	80.80	33.65	0.000	0.5787
Neutro	220	31.95	12.36	56.32	7.59	59.70	11.75	75.643	0.5051
LUC*	220	0.8	0.3	1.6	0.4	1.6	0.4	0.09	0.4341
PT	292	9.4	0.5	10.4	8.5	10.4	8.5	0.07	0.7083
APTT	292	21.62	1.87	25.30	17.95	25.00	18.33	1.773	0.4915
Blood Chemistry									
AST*	164	34.4	31.9	73.0	16.4	68.4	15.1	0.22	0.4210
ALT*	164	37.7	22.3	85.7	16.0	104.3	16.0	0.15	0.5656
ALP*	160	1230	504	2265	261	2101	260	0.0	0.9122
LDH*	152	730.6	402.5	1824.4	353.7	1619.4	348.2	0.33	0.4514
CPK*	152	310	795	931	60	1123	72	0.8	0.2041
T.Bil.*	164	0.200	0.066	0.369	0.117	0.369	0.120	0.1018	0.6001
TP	152	7.43	0.47	8.36	6.51	8.20	6.46	0.054	0.7531
ALB	144	4.11	0.28	4.66	3.55	4.70	3.46	0.030	0.6188
T.Cho.	152	115.6	24.0	163.0	68.3	173.0	72.8	59.93	0.8956
TGL*	152	39.0	17.5	88.2	17.0	79.2	17.0	0.17	0.4456
GLU*	152	83.8	13.2	113.4	60.9	110.8	61.0	0.02	0.4336
BUN	164	22.17	4.71	31.47	12.87	30.69	13.42	8.917	0.5981
CRNN	164	0.641	0.121	0.879	0.403	0.909	0.432	0.0020	0.8655
IP	152	6.052	1.001	8.030	4.074	7.576	3.696	0.3531	0.6477
Ca	152	10.14	0.54	11.20	9.08	11.24	9.24	0.134	0.5341

*) σ^2 and β were calculated for log-transformed values

Table 6: Reference intervals in female cynomolgus monkeys

				Reference Interval		Percentile		Estimates	
Parameter	**N**	**Mean**	**SD**	**UCL**	**LCL**	**97.5%**	**2.5%**	σ^2	β
Hematology									
RBC	220	5.492	0.413	6.307	4.678	6.235	4.645	0.0649	0.6197
WBC*	220	12.916	3.724	21.685	7.048	20.855	6.785	0.0214	0.7189
Ht	220	42.88	2.82	48.44	37.32	47.35	36.25	3.914	0.5083
Hb	220	13.20	0.95	15.07	11.32	15.05	10.85	0.375	0.5870
Platelet	220	450.9	109.0	665.7	236.2	674.5	259.0	2319.68	0.8046
MCV	220	78.21	3.92	85.94	70.48	86.20	70.65	2.344	0.8477
MCH	220	24.05	1.19	26.39	21.72	26.35	21.60	0.188	0.8666
MCHC	220	30.77	1.10	32.93	28.61	33.00	28.35	0.678	0.4355
Ret*	204	1.85	0.85	3.92	0.77	4.29	0.80	0.132	0.3647
Eosino*	220	2.32	1.92	8.32	0.32	7.05	0.30	0.176	0.7181
Baso*	220	0.9	0.7	3.1	0.2	2.8	0.2	0.57	0.2153
Mono*	220	3.6	1.3	7.0	2.0	6.7	2.0	0.13	0.5165
Lymph*	220	60.40	12.56	85.68	35.96	80.80	33.65	0.000	0.5787
Neutro	220	31.95	12.36	56.32	7.59	59.70	11.75	75.643	0.5051
LUC*	220	0.8	0.3	1.6	0.4	1.6	0.4	0.09	0.4341
PT	292	9.4	0.5	10.4	8.5	10.4	8.5	0.07	0.7083
APTT	292	21.62	1.87	25.30	17.95	25.00	18.33	1.773	0.4915
Blood Chemistry									
AST*	164	34.4	31.9	73.0	16.4	68.4	15.1	0.22	0.4210
ALT*	164	37.7	22.3	85.7	16.0	104.3	16.0	0.15	0.5656
ALP*	160	1230	504	2265	261	2101	260	0.0	0.9122
LDH*	152	730.6	402.5	1824.4	353.7	1619.4	348.2	0.33	0.4514
CPK*	152	310	795	931	60	1123	72	0.8	0.2041
T.Bil.*	164	0.200	0.066	0.369	0.117	0.369	0.120	0.1018	0.6001
TP	152	7.43	0.47	8.36	6.51	8.20	6.46	0.054	0.7531
ALB	144	4.11	0.28	4.66	3.55	4.70	3.46	0.030	0.6188
T.Cho.	152	115.6	24.0	163.0	68.3	173.0	72.8	59.93	0.8956
TGL*	152	39.0	17.5	88.2	17.0	79.2	17.0	0.17	0.4456
GLU*	152	83.8	13.2	113.4	60.9	110.8	61.0	0.02	0.4336
BUN	164	22.17	4.71	31.47	12.87	30.69	13.42	8.917	0.5981
CRNN	164	0.641	0.121	0.879	0.403	0.909	0.432	0.0020	0.8655
IP	152	6.052	1.001	8.030	4.074	7.576	3.696	0.3531	0.6477
Ca	152	10.14	0.54	11.20	9.08	11.24	9.24	0.134	0.5341

*) σ^2 and β were calculated for log-transformed values

2.4. Within- and Between-Animal Variations

Reference intervals for the analysis parameters shown in Tables 5 and 6 are overall reference intervals obtained from all animals judged to be healthy; however, it is thought that the range of biological variation in individual animals is much narrower than that of overall data for some parameters.

To assess within- and between-animal variations, individual within-animal variance was assumed to be similar, and analysis of variance (ANOVA) was applied using animals as a random effect.

$$y_{ij} = \mu + x_i + \varepsilon_{ij} \quad i = 1, \ldots, n; \quad j = 1, \ldots, m$$

y_{ij} represents a response variable for the j th observation of animal i. x_i are mutually independent random effects on animals, and ε_{ij} are mutually independent random errors. x_i and ε_{ij} are normal with mean 0, and variance σ_0^2 and σ^2, respectively.

Strictly speaking, σ^2 is the sum of measurement error and within-animal variance; however, it can be assumed that measurement error is insignificant when compared with within animal variance, and that σ^2 is approximately the same as within animal variance.

β is the ratio of between-animal variance to variance if defined as follows:

$$\beta = \frac{\sigma_0^2}{\sigma^2 + \sigma_0^2}$$

Estimated σ^2 and β values are also shown in Tables 5 and 6. Large β values imply

that the ratio of between-animal variance in platelets, MCV, MCH, PT, ALP, T.Cho., and CRNN in both males and females is large, and that the within-animal variations in these parameters would be much narrower in comparison with overall data variations. When we look at the relationship between mean, SD, and estimated β values average in Table 5 and 6, and not necessarily mean, SD, and β estimates are not always correlated. For example, LDH and CPK have relatively large mean and SD, but both β estimates are less than 0.5. On the other hand, CRNN and PT have relatively small mean and SD, but both β estimates are relatively high as stated before. For parameters that have large β estimates, the range of biological variation in individual animals might be much narrower than that of overall data for some parameters, thus more precise methods might be needed for drug safety evaluation.

2.5. Individual Reference Intervals

It has become increasingly common to use a minimum number of primates, including cynomolgus monkeys, in studies to assess the effects of a drug or endocrine-disrupting chemical, and it is now common to assess effects on individual animals because of the small number used. Accordingly, overall reference intervals for parameters with large β values would be too wide to be applicable and are, therefore, impractical. Using baseline

values, which are commonly obtained from untreated healthy animals in studies, the following method was proposed to calculate individual reference intervals to allow for evaluation of the effects of a drug or endocrine-disrupting chemical.

Assuming a random sample y_i of size n from a normal distribution N(θ_i , σ^2) where θ_i is from a normal distribution N(μ_0, σ_0^2). The posterior distribution of θ_i given y_i is as follows:

$$\theta i \mid yi \sim \mathrm{N}(\mu_i, \sigma_i^2)$$

where $\mu_i = \frac{n\bar{y}\sigma_0^2 + \mu_0\sigma^2}{n\sigma_0^2 + \sigma^2}$, $\sigma_i^2 = \frac{\sigma_0^2\sigma^2}{n\sigma_0^2+\sigma^2}$, and $\bar{y} = \sum_{i=1}^{n}\frac{y_i}{n}$.

The posterior distribution of unobserved y_{new} is:

$$y_{new} \mid y_i \sim \mathrm{N}(\mu_i, \sigma_i^2 + \sigma^2)$$

The 95% Bayesian predictive limit is $\mu_i \pm 1.96\sqrt{\sigma_i^2 + \sigma^2}$ (Spiegelhalter *et al*., 2004). If prior information for each animal is available, an individual reference interval using this Bayesian predictive limit can be constructed. (Koga *et al.*, 2007).

Example)

Table 7 contains data on MCH, Ht, and CPK values in four male cynomolgus monkeys in the placebo group in a toxicity study aimed at investigating effects of the drug. Data from two points before dosing and five points after dosing were obtained.

Table 7: Observed data of four male cynomolgus monkeys

				Study Day				
Parameter	Animal	Baseline1	Baseline2	Day0	Day7	Day14	Day21	Day28
MCH	1	23.2	23.4	23.6	23.0	23.7	24.1	23.7
	2	23.8	23.9	24.1	23.9	23.7	23.8	24.4
	3	24.4	24.1	24.8	25.3	25.4	24.5	25.1
	4	25.2	25.5	25.5	26.0	25.8	25.4	25.8
Ht	1	42.6	43.8	41.0	43.1	42.4	45.0	43.6
	2	42.4	45.1	43.3	44.2	44.6	46.1	45.6
	3	43.3	44.6	43.1	40.5	43.2	43.3	42.7
	4	44.3	44.5	40.1	39.7	42.2	43.3	42.7
CPK	1	74	113	174	94	86	373	536
	2	139	148	147	107	119	187	99
	3	206	115	978	350	621	236	171
	4	135	210	88	114	442	213	134

Using baseline values obtained before dosing, the individual reference interval (95% predictive limit) of each animal was estimated and is shown in Figure 6 along with the overall reference interval. Individual reference intervals are shown as dotted lines and overall reference intervals are shown as dashed lines.

As shown in Figure 6, most data after dosing were within the overall reference interval. With regard to MCH, which has a large β value (= 0.8666), individual reference intervals of all animals are much narrower than the overall reference interval, and could be a more practical index to assess individual biological variation than overall reference interval. For example, MCH data on day 14 from animal 3 was within the overall reference interval, but not within the individual reference interval of this animal. This

value is suspected to be abnormal for this animal.

The individual reference interval of CPK (β =0.2041) might be wider than the overall reference because of the small β value. As seen in the figure 6, this parameter contains a large within-animal variation, as do the estimated individual reference intervals. The individual reference interval of Ht (β =0.5083) are similar to the overall reference. For a parameter which has β =0.5 approximately, the overall reference interval might be enough for all animals.

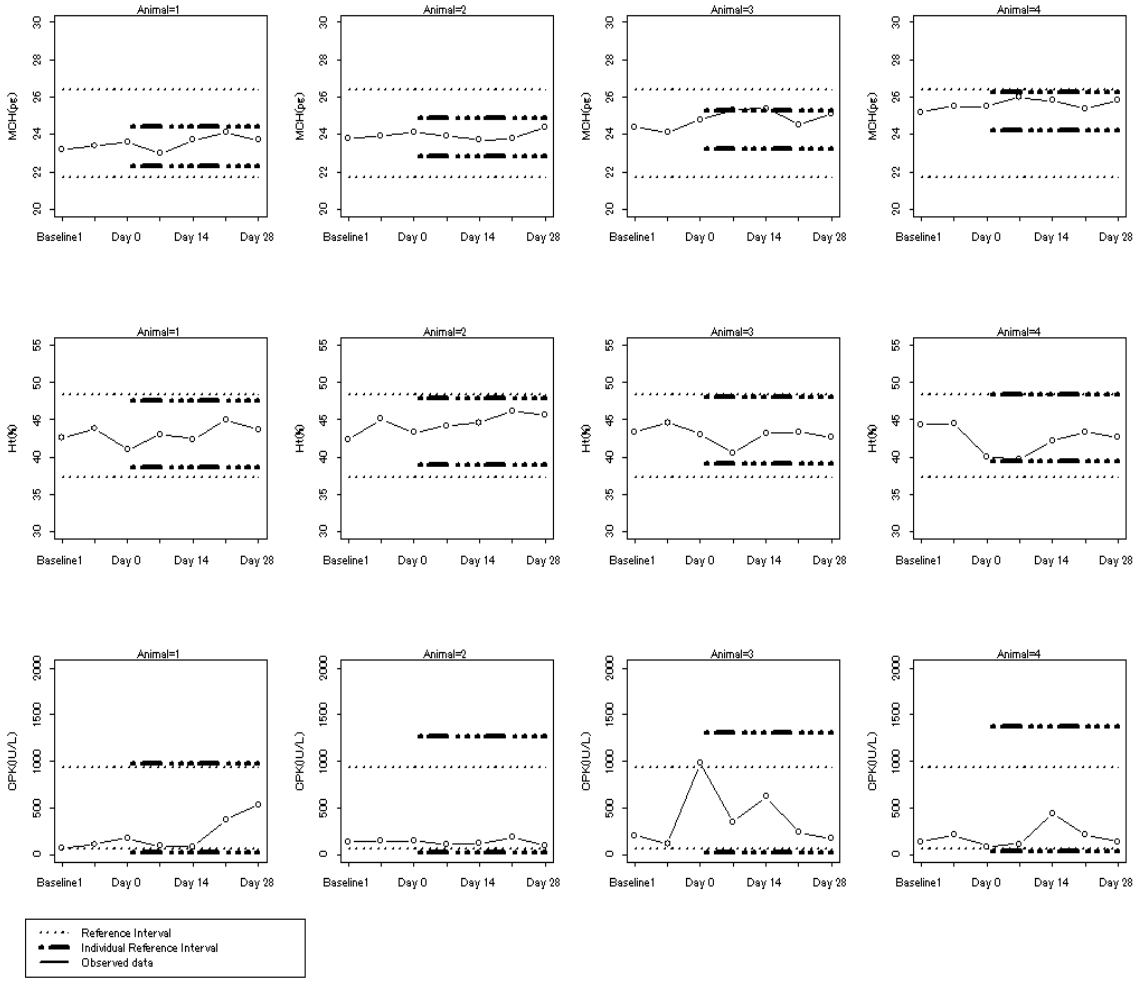

Figure 6: Individual reference intervals of four cynomolgus monkeys

2.6. Graphical Evaluation

Many kinds of statistical analysis methods have been used to evaluate health condition indices and/or drug safety based on hematological, serum biochemical, and ECG parameters. Since there are many mutually correlated parameters, complex statistical analyses should be performed theoretically. However, the results from such analyses are often confusing because the number of test parameters is so large and researchers might need time to evaluate all parameters to make a conclusion of the study.

A graphical method of readily evaluating data obtained in non-clinical studies based on background data is intuitive method for researchers who are not familiar with complex statistical analysis methods. To help researchers and to avoid frustration, a comprehensive graphical method is proposed, considering the following three steps procedure:

1. Data transformation (Box-Cox transformation)
   Transform data so that each parameter follows a normal distribution.
2. Standardization
   Transform data so that each parameter follows a standard normal distribution.
3. Radar Chart
   Show transformed data on the same chart to help researchers understand all parameter trends at a glance.

1. Data Transformation (Box-Cox transformation)

The Box and Cox (1964) transformation makes use of the parameters y^{λ} where λ is estimated from data. The parameter is defined as w, where w is given by:

$$w = \frac{y^{\lambda}-1}{\lambda} \quad \text{(if } \lambda \text{ is not equal to 0)}$$

$$= \log(y) \quad \text{(if } \lambda \text{ equals 0)}$$

The use of log(y), as the proper transformation for λ=0 is the result of the fact that the limit of (y^{λ}−1)/ λ approaches log(y) as λ approaches zero. Box and Cox showed that the profile likelihood function for λ is

$$L(\lambda) = constant - n/2 * \log(RSS(v\lambda))$$

where $v\lambda = \frac{y\lambda}{y^{\lambda-1}}$, y is the geometric mean of the observations, and $RSS(v\lambda)$ is the residual sum of squares for the regression of $v\lambda$.

Box and Cox (1964) suggested using the profile likelihood function for the largest linear model to be considered as a guide in choosing a value for λ, which will then remain fixed for any remaining analyses. To find an appropriate λ for each parameter, the boxcox function in S-Plus, which calculates and displays the Box-Cox profile likelihood function, is used.

2. Standardization

If an ideal λ for each parameter that follows a distribution of a certain family can be found, each transformed parameter, w, would follow a normal distribution approximately. The central limit theorem justifies the use of the normal distribution in many applications. The density function of the normal distribution depends on two parameter, μ and σ, where

$-\infty < \mu < \infty$, $\sigma > 0$.

$$f(w) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(w-\mu)^2}{2\sigma^2}\right)$$

Consider the following transformation:

$$z = \frac{w - \mu}{\sigma}$$

It can then be shown that z follows a standard normal distribution, with $\mu = 0$ and $\sigma = 1$. A confidence interval for the population mean, μ, is now considered. It is well known that the 95% confidence interval for mean is -1.96 to 1.96, or approximately -2 to 2.

3. Radar Chart

Radar charts are often used to show the relative frequency of data measures in quality control problems. On a radar chart, statistics are displayed along spokes that radiate from the center. The charts are stacked on top of one another with reference circles, thus giving them the appearance of a radar screen. Usually, the chart vertices, the points where the statistical values intersect then spokes, are based on the frequencies associated with the levels of a single numeric variable, but standard values that approximately follow the standard normal distribution are used here.

Liu *et al*. (2008) used radar charts to show clinical pathology data from cynomolgus

monkeys from China in which diarrhea was observed during quarantine. They imported and quarantined 3,148 cynomolgus monkeys (2.5 to 6.5 years) from China to Japan. They obtained the hematology and blood biochemistry data form these monkeys on Day 32 of quarantine and analyzed separately by sex. There were 2,890 animals in which no abnormalities were observed during the 35-day quarantine period (healthy group), and 258 animals which exhibited diarrhea 1 to 12 times (diarrhea group). They followed the procedure above, and analyzed the clinical pathology data from 11 animals exhibiting diarrhea repeatedly. Using radar charts, they've found that there were significant differences in PLT and ALP in both sexes. See figure 7 as an example of the radar chart.

.

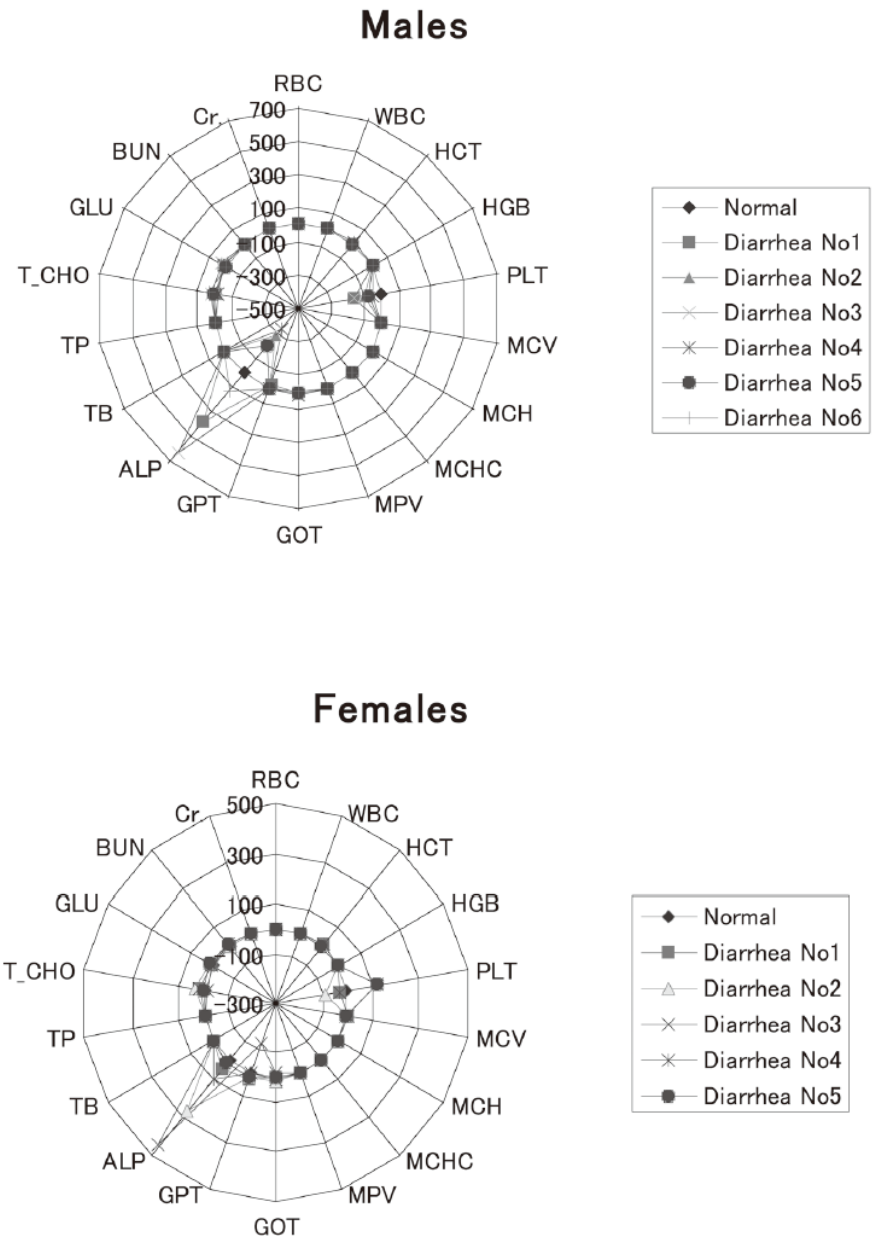

Figure 7: Example of radar chart (Liu *et al*., 2008)

2.7. Comments

Summary statistics and histograms reveal the distribution of hematological and serum biochemical parameters in cynomolgus monkeys. Many parameters did not follow normal distribution, and this should be taken into account when performing statistical tests such as the t test and analysis of variance (ANOVA), which assume that data follow normal distribution. And we also need to understand that there might be a parameter that is difficult to assume well-known statistical distribution like ALP (Alkaline phosphatase) which is used to help detect liver disease or bone disorders.

It is essential to use accurate reference intervals to assess health condition, biological variation, or effects of a drug, or endocrine-disrupting chemical on the study system. For this purpose, reference intervals for hematological and serum biochemical parameters in cynomolgus monkeys are shown. Within- and between-animal variations were also evaluated, and unlike other animals such as rats and beagle dogs, it was found that some parameters showed large between-animal variations in cynomolgus monkeys. This would be natural because this primate is close phylogenetically relationships to humans. Among hematological and serum biochemical parameters, between-animal variations were large in platelet, MCV, MCH, PT, ALP, T.Cho., and CRNN in both males and females. Thus, the within-animal variations in these parameters would be much narrower in comparison

with overall data variations. As shown in the example, not only overall reference intervals but individual reference intervals calculated from only two baseline values can yield useful information for evaluation of biological variation for these parameters. More baseline information might be necessary; however, individual reference intervals of these parameters for each animal estimated by the proposed method would make it possible to assess biological variations or effects of a drug or endocrine-disrupting chemical more accurately.

As we proposed, a graphical method would be useful to evaluate safety profiles in clinical and non-clinical studies when the background data is available. Usually, there are many parameters to evaluate drug safety profiles. We believe that it is a comprehensive and intuitive method that can make researchers understood a whole data at a glance.

It has been reported that some parameters in human beings change with age (Wright and Royston, 1997; Royston, 1991; Tango, 1998). And, various relationships between biological parameters, adult stature, and timing of the pubertal growth spurt also has been investigated in humans (Sumiya *et al*., 2001; Qin *et al*., 1996). In cynomolgus monkeys, parameters such as ALP, RBC, and body weight are also considered to vary with age (Yoshida *et al*., 1986). Thus, data continuously obtained from the same animals over a long period are necessary to estimate appropriate age-related reference intervals for these

parameters.

We also realized that the background data and reference intervals in cynomolgus monkeys are affected by analytical methods and animal maintenance conditions. These data should be updated appropriately to have more precise information.

3. Safety Pharmacology Study

As stated before, the QT interval is considered identified as a safety biomarker for the assessment of the risk of ventricular tachycardia. Ventricular depolarization usually found a small component of the QT interval. However, ventricular repolarization time will extensively shorten as heart rate increases, and extensively lengthen as heart rate decreases. Currently, it is considered important to correct the QT interval for the HR or RR interval to compare QT intervals that were measured at different HR/RR intervals. To investigate appropriate QT correction method in cynomolgus monkeys, we conducted a non-clinical study with twelve male cynomolgus monkeys.

The differences between a thorough QT study and a safety pharmacology study are explained briefly, and then, QT correction methods in cynomolgus monkeys are discussed.

3.1. Study Design

In 2005, ICH introduced the concept of new study, called a thorough QT study in ICH E-14 guideline; “The Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs”.

The primary purpose of a ‘thorough QT/QTc study’ is to determine if an experimental drug has an effect on cardiac ventricular repolarization, which means to evaluate changes

in QT or QTc interval during drug treatment. Other purpose of a thorough QT/QTc is to demonstrate that the study is capable to detect differences of regulatory concern (around 5msec, practically). Therefore, the thorough QT/QTc will be designed to confirm that any lack of detected change is because of actual lack of change rather than lack of assay sensitivity of the study.

The thorough QT/QTc is generally carried out in healthy volunteers to determine if an experimental drug has an effect on the QT/QTc interval in target patient populations. Volunteers will be screened for normal cardiac electrical activity for precise measurement of the ECG and to control confounding factors. Current ICH E 14 Guidance requests all sponsors submitting new drug applications to conduct a thorough QT/QTc. Generally, a thorough QT/QTc will be conducted in early clinical development after some knowledge of the pharmacokinetics of the drug.

ICH E-14 guideline requires that the thorough QT/QTc should be adequate and well-controlled, with deal with potential bias, including use of randomization, appropriate blinding, and concurrent placebo control group. It is considered to be important to have a high degree of confidence in the ability of the thorough QT/QTc to detect differences of clinical significance. Currently, this confidence in the ability of the thorough QT/QTc to detect differences of clinical significance is considered to be determined by the use of a

positive control to confirm assay sensitivity.

The thorough QT/QTc could be a cross-over design when pharmacokinetic information of the experimental drug is obtained sufficiently. There are some cases that a parallel group design would be preferred. If the experimental drug or its metabolites have long elimination half-lives, or they require relatively long treatment interval, the parallel group design is preferred. Or if carry-over effects are noticeable, for example, long-lived active metabolites, the parallel group design is also preferred.

The choice of the study design (parallel versus cross over) is a usually straightforward consideration. Practically, the cross over design is often selected because it has some advantages compared to the parallel group design. Usually, in comparison with the parallel group design, the cross-over design needs smaller sample size of subjects. In the cross-over design, subjects would have their own control data and we can reduce variability of differences between placebo and the experimental drug related to inter-subject variability by using appropriate statistical analysis method. Also, in the cross-over design, we might be able to find QT interval correction method based on individual data.

Generally, the thorough QT/QTc is a four-arm study. The typical treatment groups are:

- Active control drug (typically oral moxifloxacin) to establish assay sensitivity.
- Placebo, for comparison of response to doses of the experimental drug.

- Proposed therapeutic dose of the drug.
- Supratherapeutic dose, several multiples of the therapeutic doses, to mimic the worst case scenario in patients if the drug were to be approved.

The administrations of the experimental drug doses and placebo are double-blind, and ECG measurements and readings are also performed blinded to associated treatments, subjects, and time of ECG.

A positive control is currently included in a thorough QT/QTc to assess the sensitivity of the study. A positive control is expected to have an effect on QT interval in range of 5 to 10 msec. Orally administered moxifloxacin 400mg is the most commonly used positive control in thorough QT/QTc studies. Pharmacokinetic aspects and the effect on QT interval of moxifloxacin have been investigated by many researchers (Demolis *et al*., 2000).

In the cross over design, each subject will be randomly assigned to receive one of a set of groups of treatments within the study. Each group represents a pre-specified administration order of the study treatments. Between treatments period, there would be sufficient wash-out periods (approximately 5 times of the maximum half-life time of the experimental drug or positive control).

Table 8: Williams Design

	Period 1	Period 2	Period 3	Period 4
Group 1	Treatment	Supra	Placebo	Control
Group 2	Supra	Control	Treatment	Placebo
Group 3	Placebo	Treatment	Control	Supra
Group 4	Control	Placebo	Supra	Treatment

Treatment: Therapeutic dose, Supra: Supratheraputic dose, Placebo: Placebo, Control: Active control

As an example, suppose a cross over study has four treatments and four periods. The four groups can be expressed as in Table 8.

In this design, each treatment will be equally precisely compared with every other treatment (Jones and Kenward, 2003). It is recommended to use a balanced Williams design (Williams, 1949). Note that the example has even number of treatments, thus a square design can be used. If a study has odds number of treatments, we need two square designs and the number of groups is 2 times larger than the number of treatments.

If the statistical analysis model includes period, sequence, treatment, first-order carryover, and direct-by-carryover interaction as fixed effects, the design based on one Williams square, as shown in Table 8, does not have sufficient degrees of freedom to assess direct-by-carryover interaction (Chen and Tsong, 2007).

Table 9: Two Latin Square Design

	Group	Period 1	Period 2	Period 3	Period 4
1	Group 1	Treatment	Placebo	Control	Supra
	Group 2	Supra	Control	Placebo	Treatment
	Group 3	Placebo	Treatment	Supra	Control
	Group 4	Control	Supra	Treatment	Placebo
2	Group 5	Treatment	Supra	Placebo	Control
	Group 6	Supra	Treatment	Control	Placebo
	Group 7	Placebo	Control	Treatment	Supra
	Group 8	Control	Placebo	Supra	Treatment

Treatment: Therapeutic dose, Supra: Supratheraputic dose, Placebo: Placebo, Control: Active control

An alternative design is to use multiple Latin square designs. A Latin square design requires that all four treatments appear in each row and in each group once. And, balanced Latin square design requires the number of one treatment before other treatments is equal to the number of other treatments before one treatment. A Williams square design is more preventive than a Latin square design. When we can assume that there is no carryover effect, Latin square is the most efficient crossover design. In order to improve the degrees of freedom in assessing direct-by-carryover interaction, we can select multiple distinct Latin square design such as shown Table 9.

On the other hand, a typical design of safety pharmacology study, is a dose escalation design, or often called as a step-up design as shown Table 10, even though we don't know if we can assume there is no carryover effect.

Table 10: Example of step-up design

Animal	Period 1	Period 2	Period 3	Period 4
1	Placebo	Low	Middle	High
2	Placebo	Low	Middle	High
3	Placebo	Low	Middle	High
4	Placebo	Low	Middle	High

High: High dose, Middle: Middle dose, Low: Low dose, Placebo: Placebo

This is because dose levels in safety pharmacology studies are usually much higher than those of clinical studies. We need to assess a drug's potential to delay ventricular depolarization in human beings, a risk-benefit assessment in order to select a suitable candidate drug for clinical development based on efficacy, safety, and pharmacology properties prior to the first clinical study in humans.

3.2. Statistical Analysis Model

Although many statistical analysis models are proposed, if the study design is the Williams design as shown Table 8, a typical statistical analysis model would include period, sequence, treatment, as fixed effects. Let y_{ijkl} be an endpoint related to QT interval of the k-th recording time point of the j-th subject in the l-th sequence who receives the i-th treatment, where i = treatment, supra, placebo, control; $j = 1, \ldots, n$; $t = 1, \ldots, k$; and $l = 1, \ldots, 4$. For any given k-th time point, an example of the model of this crossover

design at the *t*-th time point is

$$y_{ijtl} = \mu_{it} + \beta_{itl} + \gamma_{ijt} + \varepsilon_{ijtl}$$

μ_{it} - the mean of *i*-th treatment at the *t*-th time point,

β_{itl} - the mean of period effect ; it can also be used as a random effect with mean 0 and variance σ^2_{tp}

γ_{ijt} - independent and identically normally distributed random subject effects with mean 0 and variance σ^2_{ts},

ε_{ijtl} - independent and identically distributed normal random errors with mean 0 and variance $\sigma^2_{t.}$.

Note that $\sigma^2_{tp} = \sigma^2_{p}$, $\sigma^2_{ts} = \sigma^2_{s}$ and $\sigma^2_{t.} = \sigma^2$, when the variance of β_{itl} , γ_{ijt}, and ε_{ijtl} are equal across all time points.

Similarly, if the study design is the step-up design as shown Table 10, a typical statistical analysis model would include only treatment, as a fixed effect assuming no carry over effect. Let y_{ijk} be an endpoint related to QT interval of the *k*-th recording time point of the *j*-th subject in the *i*-th treatment, where *i* = Treatment, Supra, Placebo, Control; *j* = 1, …, *n;* and *t* = 1, …, *k.* For any given *k*-th time point, an example of the model of this crossover design at the *t*-th time point is

$$y_{ijt} = \mu_{it} + \gamma_{ijt} + \varepsilon_{ijt}$$

μ_{it} - the mean of *i*-th treatment at the *t*-th time point,

γ_{ijt} - independent and identically normally distributed random subject

effects with mean 0 and variance σ^2_{ts},

ε_{ijt} - independent and identically distributed normal random errors with mean 0 and variance $\sigma^2_{t.}$.

Note that $\sigma^2_{ts} = \sigma^2_{s}$ and $\sigma^2_{t.} = \sigma^2$. when the variance of γ_{ijt}, and ε_{ijt} are equal across all time points.

3.3. ECG Time Points

In the thorough QT/QTc, ECG is often measured one or more baseline time points before experimental drug is administrated. And also, on the day before the first treatment administration, ECGs are measured at all time-points corresponding to those at which ECGs will be measured after the experimental drug administration.

In a safety pharmacology study, ECG data from the placebo group are often used to assess effect of treatment at each time point. The time points of the ECG measurement after administration should be determined using the pharmacokinetic information of the experimental drug. ECGs are usually measured as a set of replicates. Practically, 3 ECGs are often measured at 1 minute interval of 10 seconds in duration.

3.4. QT Interval Correction Formulae

As stated before, QT interval is representative of ventricular depolarization

/repolarization time. The prolongation of QT interval is associated with serious cardiac malfunction, including TdP and sudden cardiac death. Accordingly, systematic evaluation of any QT interval prolongation induced by new chemical entities is required at both the non-clinical and clinical stages of the drug development process.

QT interval is inversely related to heart rate. Typically, QT interval increases with increasing RR interval in the electrocardiogram (ECG). Many formulas have been applied to correct QT interval from RR interval change artifacts. The best known of these are Bazett's (1920) formula and Fridericia's (1920) formula, which are used in both clinical studies and non-clinical studies. These formulas are simple and convenient, but Malik (2001), Desai *et al*. (2003), and Molnar *et al*. (1996) have reported limitations in their application to the clinical development process.

However, the ICH E-14 guideline recommends to submit QT, RR, HR intervals, as well as QT interval corrected using Bazett's and Friericia's corrections, therefore, QTc intervals corrected using these formula are still commonly used in both clinical studies and non-clinical studies for practical reasons.

Bazett's Correction Formula:

$$QTcf = \frac{QT\sqrt{1000}}{\sqrt{RR}}$$

Fridericia's Correction Formula:

$$QTcf = \frac{QT\sqrt[3]{1000}}{\sqrt[3]{1000}}$$

Several study population QT correction methods are becoming popular in both clinical studies and non-clinical studies. Two of these methods are linear and log-linear regression model. These methods use pooled ECG readings measured before any treatments.

Linear regression:

$$QT = a + bRR$$

$$QTc = QT + b(1 - RR)$$

Log linear regression:

$$\log(QT) = c + d\log(RR)$$

$$QTc = \frac{QT\sqrt[d]{1000}}{\sqrt[d]{RR}}$$

The correction methods have some statistical assumptions and might perform poorly if assumptions are not met.

Some formulas to correct QT interval from RR interval change artifacts in rhesus monkeys and beagle dogs (Hashimoto *et al.*, 2002; Raunig *et al.*, 2001; Miyazaki and Tagawa, 2002) have been investigated. Although the cynomolgus monkey is one of the closest experimental animal species to humans, and commonly used not only in toxicological, pharmacological, and cardiovascular researches, only a few formulas for

correcting QT interval in this species have been investigated (Gauvin *et al.*, 2006; Hayes *et al.*, 1994).

3.5. Telemetry ECG Study in Cynomolgus Monkeys

Unlike well-known uniform correction methods, individual QT correction methods for humans have been developed in tandem with the development of statistical analysis software. We conducted a non-clinical study to determine which correction formula is most desirable to cynomolgus monkeys. The study design is summarized below.

3.5.1. Study System

All procedures involving animals were approved by the Animal Care and Use Committee of Shin Nippon Biomedical Laboratories, Ltd. (Kagoshima, Japan), and were performed in accordance with the standards published by the National Research Council (Guide for the Care and Use of Laboratory Animals, NIH OACU) and the National Institute of Health Policy on Human Care and Use of Laboratory Animals. All work was performed at Shin Nippon Biomedical Laboratories, Ltd. (Kagoshima, Japan).

Twelve male cynomolgus monkeys (aged 4–6 years, weighing 3–6 kg) were housed

in a controlled environment maintained at a temperature of between 23 °C and 29 °C and relative humidity of between 35% and 75%. The animal room was ventilated with a minimum of 15 air changes/h, and a light/dark cycle of 12 h (lights on at 08:00) was set.

3.5.2. Surgical Implantation

A Data Sciences International (DSI) telemetry transmitter (TL11M2-D70-PCT, Data Sciences International Inc.) was implanted intraperitoneally, and fixed inside the abdominal wall under ketamine hydrochloride (10 mg/kg, intramuscular, Fuji Chemical Industry Co., Ltd.) anesthesia. Two unipolar lead electrodes for ECG were implanted subcutaneously at predetermined locations on the right manubrial border of the sternum and left anterior auxiliary line sixth rib. An antibiotic [aqueous suspended injection of dihydrostreptomycin sulfate and benzyl penicillin procaine: dihydrostreptomycin sulfate (250 mg potency/mL), benzyl penicillin procaine (200,000 units/mL)] was administered intramuscularly at 0.05 mL/kg to the animals once daily for three days including the day of implantation (Horii *et al*., 2002). The animals were allowed approximately two or three weeks to recover from surgery and were used in the present study after the ECG parameters stabilized.

3.5.3. ECG Measurement

QT and RR intervals in monkeys in which Data Sciences radio-telemetry devices had been implanted were obtained using ECG Processor (SBP-2000, Softron Co., Ltd.) data analysis and acquisition software. ECG measurements were recorded for 24 hour and each interval was calculated from averaged continuous ECG waveforms during an eight-second period every 30 min.

3.6. Individual QT Correction Formula

After examining QT and RR interval distributions, we selected a linear model for log-transformed QT and RR intervals, similar to the model evaluated by Shah and Hajian (2003) and Li *et al.* (2004). A similar approach has been adopted by Malik (2001), Miyazaki and Tagawa (2002), and Gauvin *et al*. (2006), but as Shah and Hajian (2002) have remarked, that described above has the advantage of enabling maximum likelihood estimation (Shah and Hajian, 2003). Since mean QT and RR intervals are considered to shift between day and night (Miyazaki and Tagawa, 2002; Gauvin *et al*., 2005), we also included the effect of light/dark cycles in the model.

Model 1:

$$\log(QT_{ij}) = \alpha_i - \beta_i * \log(RR_{ij}) + \varepsilon_{ij}$$
$$\alpha_i = \alpha + a_i, \qquad \beta_i = \beta + b_i$$

Model 2:

$$\log(QT_{ij}) = \alpha_i - \beta_i * \log(RR_{ij}) + \gamma * I[j] + \varepsilon_{ij}$$
$$\alpha_i = \alpha + a_i, \qquad \beta_i = \beta + b_i$$
$$I[j] = 1 \text{ (if j is the lighting period)},$$
$$= 0 \text{ (otherwise)}$$

In both models, *ij* represents the *j*-th observation from animal *i*, and ε_{ij} are *iid* random errors. β_i is the individual correction factor for animal *i* and varies about the mean correction factor β. α_i is a nuisance factor for animal *i* and varies about the mean correction factor α. In model (2), I[j] is an index function and γ represents light/dark effect. For estimation, we assume that the random effect, (a_i, b_i), *i.i.d.* N(0, *V*), where *V* is a 2 by 2 arbitrary covariance matrix, and ε_{ij} is *i.i.d.* N(0, σ^2)

The correction formula for Model 1 is:

$$QTc1_{ij} = QT_{ij} * \left(\frac{1000}{RR_{ij}}\right)^{\beta_i}$$

The correction formula for Model 2 is:

$$QTc2_{ij} = QT_{ij} * \left(\frac{1000}{RR_{ij}}\right)^{\beta_i} / e^{\gamma * I[j]}$$

Koga *et al.* (2007).

The linear models were fitted using the "proc mixed" procedure of SAS (SAS Institute, Cary, NC USA, 1999). QT values corrected by Bazett's and Fridericia's formulas are presented for comparison purpose. Note that these correction formulas are equivalent if we assume that in $QTc1_{ij}$, β_{ij}=1/2 or 1/3.

Bazett's and Fridericia's formulas assume a constant relationship between QT and RR intervals. On the other hand, baseline/placebo data are required for all animals when individual correction formulas, like Model 1 and Model 2, are used to evaluate the cardiac safety of drug candidates. Accordingly, the number of baseline/placebo data required for implementing an animal-specific or individual correction formula for each animal was investigated. Sub-sampling simulation was performed for each sample size in which the number of repeat measurements of QT-RR data for each animal varied between 3 and 42. The simulation was repeated 1000 times for each sample size. It was assumed that the values for α_i and β_i, obtained by using complete data, were "true" values for each animal, and the following index was used to determine the sample size required to estimate α_i and β_i.

$$MSE_n = \frac{1}{m}\sum_i (\beta_{in} - \beta_i)^2 \quad i = 1, \ldots, 12, \quad n = 3, \ldots, 42$$

3.7. Result from the Telemetry ECG Study in Cynomolgus Monkeys

Results from the telemetry EGC study are presented in this section. Some basic information is described first, and then results using Model 1 and Model 2 introduced in the previous section.

3.7.1. Graphs

Histograms of QT and RR intervals are shown in Figure 8. These histograms and non-parametric density lines show that both QT and RR intervals had right-heavy tail distributions, suggesting that it cannot be assumed that these parameters follow normal (gaussian) distributions.

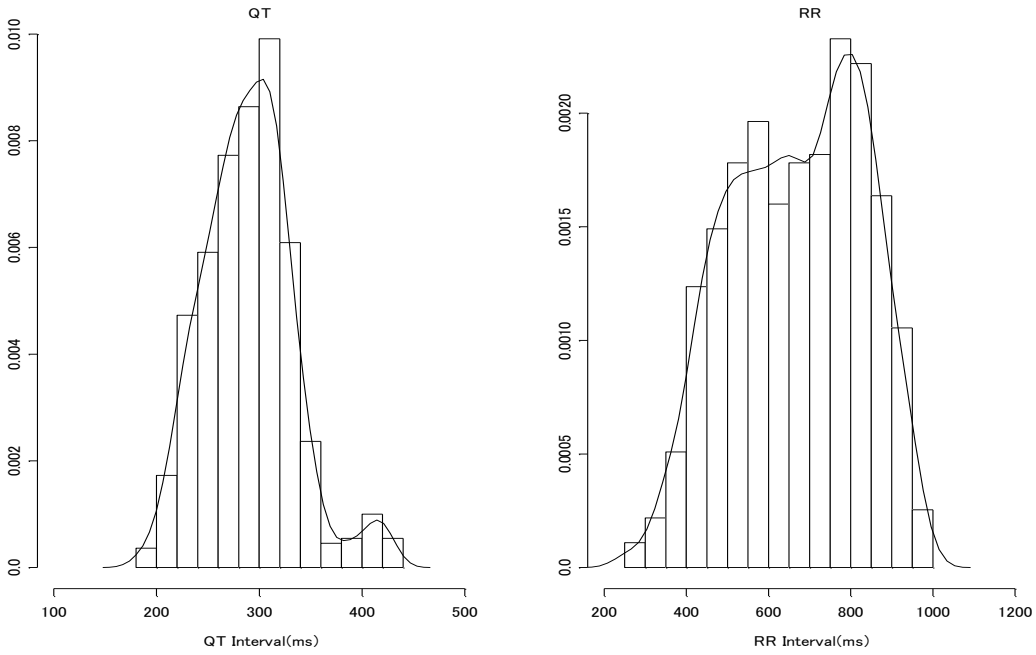

Figure 8: Histograms of QT and RR intervals for 12 male cynomolgus monkeys (N=550)

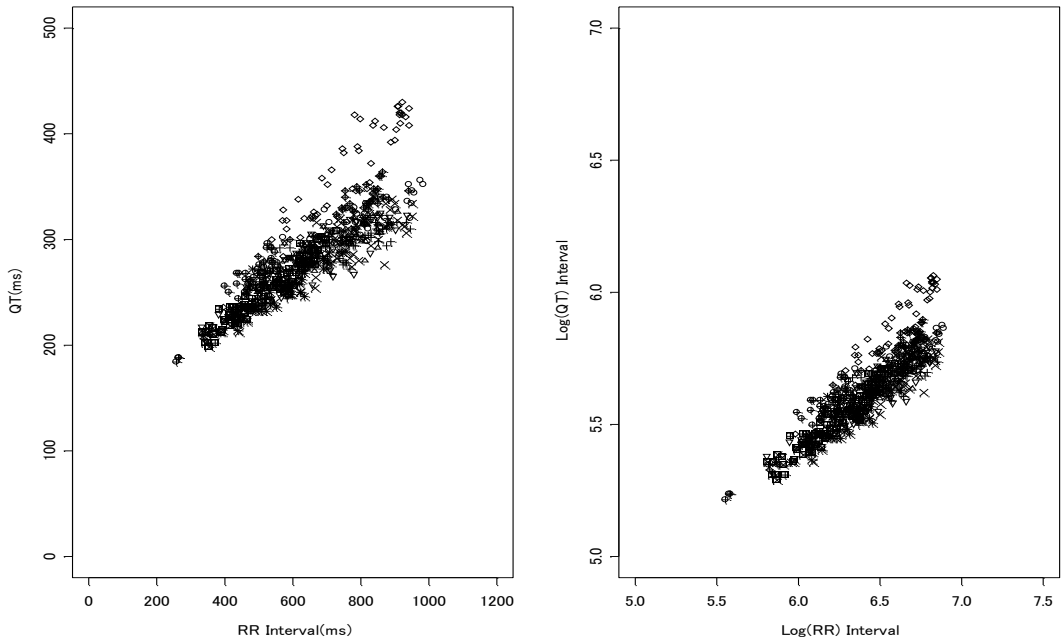

Figure 9: Scatter plots of QT interval versus RR interval, and log (QT) versus log (RR) for 12 male cynomolgus monkeys (N=550)

Figure 9 shows the correlation between QT and RR intervals. The correlation coefficient for the QT and RR intervals was 0.8381, and the correlation coefficient for log-transformed QT and RR intervals was 0.8660. Figure 9 also shows that data from a given animal are distributed over a relatively narrow range. This implies models including Model 1 and Model 2 in the previous section based on log-transformed QT and RR intervals, are more preferable than simple linear models.

3.7.2. QT Correction Formula

An estimate of β=0.4868 with b_i ranging from −0.0578 to 0.1607 was obtained from the Model 1; thus the individual correction factor β_i ranged from 0.4290 to 0.6475. Similarly, estimates of β=0.4233 with b_i ranging from −0.06589 to 0.17103 and of γ=−0.04864 were obtained from the Model 2; thus the individual correction factor β_i ranged from 0.3662 to 0.5944.

Figure 10 shows four types of corrected QT values (QTcb, QTcf, QTc1i and QTc2i) with regression lines for all animals. Graphically, it can be seen that the regression lines for QTc1i and QTc2i were closer to the horizontal than those for QTcb and QTcf. Bias associated with QTcb and QTcf, corrected by Bazett's formula and Fridericia's formula was also found.

Table 11 shows estimated within- and between-animal variability for these four types of corrected QT, and among these, QTc2i showed the least within-animal variability and QTcf showed the greatest within-animal variability.

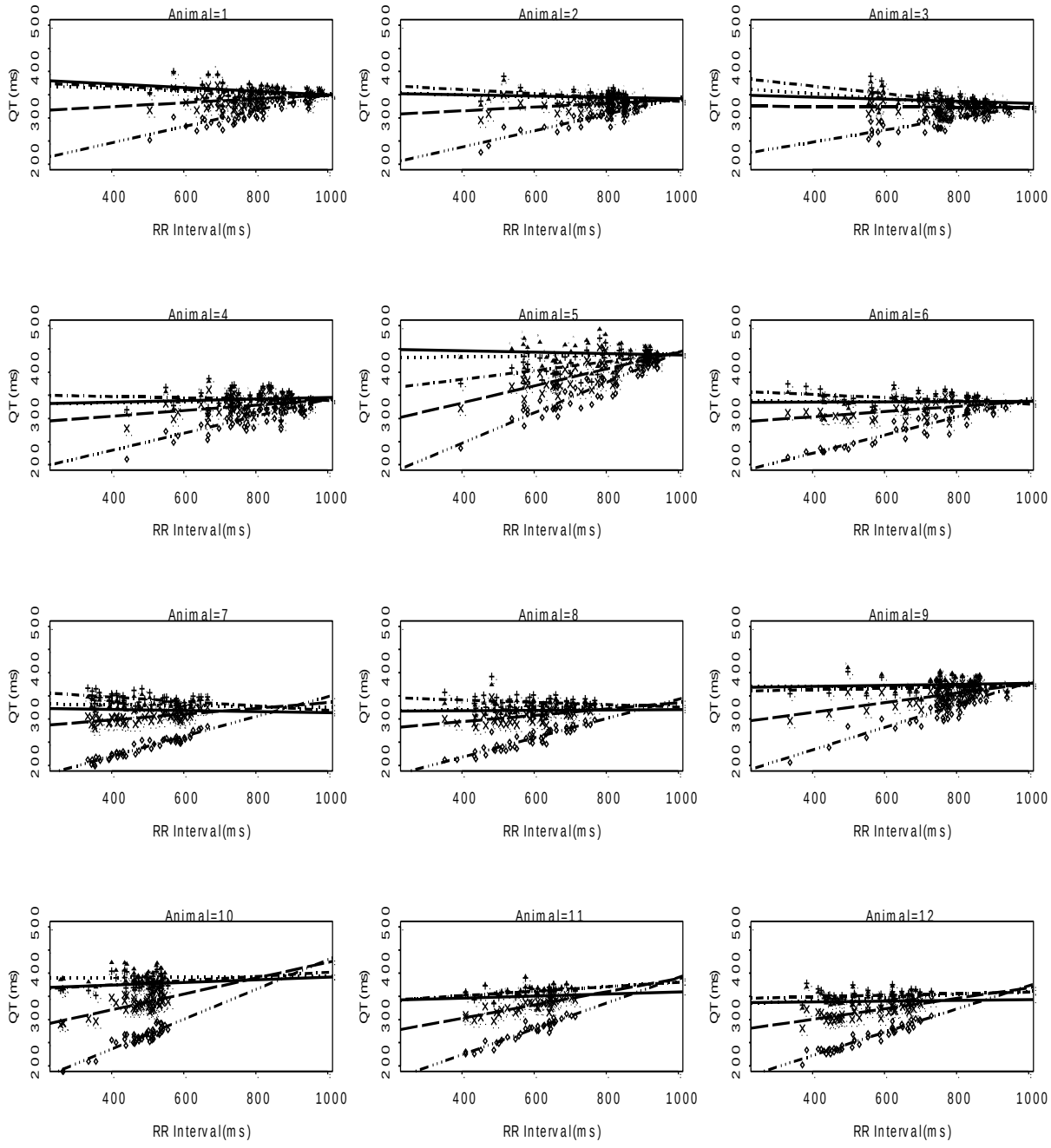

Figure 10: Scatter plots of QT interval corrected by four different formulae, and the original QT intervals versus RR interval for all animals.

Lines represent the regression line fits for QTcb (Bazett's formula), QTcf (Fridericia's formula), QTc1i (Linear model (1)), and QTc2i (Linear model (2)).

Table 11: Mean sum of squares of within- and between-animal variability of four corrected QT intervals

QT	Within-animal variation	Between-animal variation
Bazett's formula QTcb	264.1	26302.6
Fridericia's formula QTcf	322.2	30558.4
Linear model (1) QTc1i	238.7	49944.1
Linear model (2) QTc2i	170.8	53358.8

3.7.3. Sample Size Estimation

Figure 11 shows the results of simulation carried out to determine the number of observations required to estimate values of β_i accurately. Visually, MSE values were very high when n was less than or equal to 10 and almost identical when n was 24 or greater for both QTc1i and QTc2i. From this result, we need 24 or more baseline data for each animal to estimate individual correction coefficient, β_i, accurately.

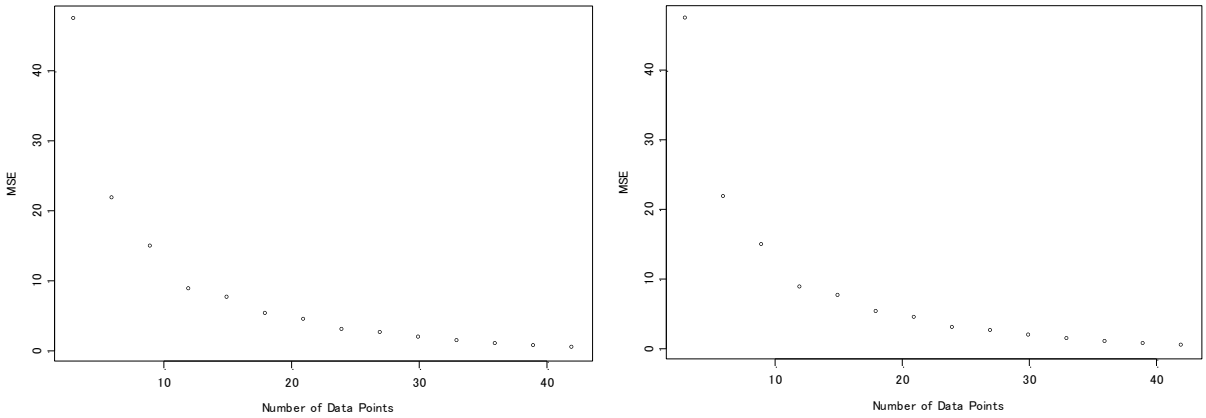

Figure 11: Mean of MSE. 1000 observations were simulated for each number of data points.
(left: Linear model 1, right: Linear model 2)

3.8. Comments

From the results of our telemetry ECG study, histograms and non-parametric density lines of QT and RR intervals in cynomolgus monkeys show that the distributions of these parameters had right-heavy tail distributions. This suggests that it cannot readily be assumed that these parameters follow the normal (Gaussian) distribution, and it is considered that this should be taken into account when performing statistical analyses. For this reason, the linear model for log-transformed QT and RR intervals was chosen to correct the QT intervals for the effect of RR interval in cynomolgus monkeys.

From scatter plots, it was found that the data from a given animal were distributed across a relatively narrow range, suggesting that a constant relationship between QT and RR intervals in cynomolgus monkeys cannot readily be assumed, and that an animal-

specific model is required to correct QT interval values in individual animals.

The estimated values obtained with the linear model also strongly support the need for an animal-specific model. These estimates (β was 0.4868 and β_i ranged from 0.4290 to 0.6475 with Model 1, and β was 0.4233 and β_i ranged from 0.3662 to 0.5944 with Model 2) suggest that each animal requires an individual correction factor varying about the mean correction factor β. Therefore, as in humans, it is considered that a constant relationship between QT and RR intervals in cynomolgus monkeys cannot be assumed.

By assessing within- and between-animal variability, we found that QTc2i had the lowest within-animal variability, representing a desirable aspect from which to maintain a constant QT value for a given RR interval change and to facilitate detection of small changes in QT interval. It was also confirmed from the scatter plots that QTc2i was the most appropriate method for maintaining a constant QT value. The regression line fits for QTc2i for each representative animal were closer to the horizontal than those for QTcb and QTcf. Moreover, it was found that between-animal variability was greater than within-animal variability, meaning that QT interval varies widely between individual cynomolgus monkeys. Again, it is considered that a constant relationship between QT and RR intervals in cynomolgus monkeys cannot readily be assumed.

We believe that an individual correction method, such as that proposed here, is

preferable to a uniform correction formula such as Bazett's or Fridericia's for evaluating QT intervals in cynomolgus monkeys. And as we have learned that QTc2i shows the least within-animal variation, we concluded that where possible, an individual correction method on the basis of the circadian rhythm is more preferable.

However, even if an appropriate QT correction formula is chosen, it will not always be possible to obtain satisfactorily corrected QT values. The models used here require extensive baseline/placebo data for all animals to determine the individual correction factor β_i. Commonly, only a few baseline data are collected in a safety pharmacology study, but the simulation results shown here indicate that at least 24 pairs of QT-RR (repeat measurement) data are required for a given animal to obtain a stable individual correction factor for that animal. Ideally, these pairs of QT-RR baseline data would be measured in animals under normal conditions over a wide range of RR values to calculate β_i more accurately.

If sufficient baseline data are not available, Bazett's formula could be considered as an alternative method for measuring QT interval in cynomolgus monkeys since the value for β obtained from the linear models were 0.4868, and 0.4233, which is relatively close to Bazett's correction factor of β=0.5. This may explain the popularity of Bazett's formula in empirical calculations with data obtained from cynomolgus monkeys. This simple

formula might be sufficient to reveal average QT interval change, but it is considered unsuitable for accurately evaluating QT interval prolongation.

It was concluded that an individual correction method is recommended for evaluating QT interval from RR interval change artifacts in cynomolgus monkeys (Koga *et al*., 2007). It is still unknown how accurately we can detect drug-induced QT interval prolongation.

4. Discussion

Torsade de Pointes (TdP) is a specific, potentially fatal arrhythmia associated with prolongation of the QT interval on the surface electrocardiogram (ECG). The arrhythmia may present with no symptoms, or with palpitations, dizziness, syncope, or sudden death. As awareness that drug induced QT prolongation, non-cardiovascular drugs might cause QT prolongation, the ICH issued the guideline E14: The Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non Antiarrhythmic Drugs. Currently, this guidance requires an ECG clinical study, commonly known as a thorough QT/QTc, which is conducted with new available drugs. The thorough QT/QTc is designed to exclude threshold drug effect below which QT interval changes are considered to have no significant clinical concern. (A “negative” thorough QT/QTc study). Since ICH E14 guideline was issued, many thorough QT/QTc studies have been conducted and wide-ranging experience shows the sensitivity of the thorough QT/QTc. On the other hand, the thorough QT/QTc has been criticized for its low cost effectiveness for timeline and lack of specificity in the pharmaceuticals industry. (Bouvy *et al*., 2012).

The non-clinical assessment to determine the potential of drugs to prolong cardiac repolarization is usually determined by non-clinical studies stated in the ICH S7B guideline. Assessment of drug effects on the human ether-a-go-go-related gene (hERG)

potassium channel assay and safety pharmacology studies in a nonrodent (usually dog or monkey) are primarily designed to detect relatively large effects that might be a concern in the clinical studies. Although there is large variability in the conduct of the non-clinical studies, it is reported that these studies have good predictiveness for QT prolongation in human beings. (Ewart *et al*., 2014; Koerner *et al*., 2013). However, the possibility of non-clinical studies to predict the result of a thorough QT/QTc in human is still questionable. One of reasons is that regulatory authority's concern is relatively small magnitude of the drug effect on QT intervals which is considered not to detect in non-clinical studies.

In section 2, we conduct a non-clinical study using cynomolgus monkeys to obtain background data of this primate. Fundamental information of various hematological and serum biochemical parameters in cynomolgus monkeys, often used for a safety pharmacology studies, are shown. Also, methods to use background information and reference intervals were introduced to evaluate drug safety more precisely. It was found that many parameters don't follow normal distributions, which are often assumed when we perform statistical analysis. It also revealed that some parameters showed large between-animal variations in comparison with within-animal variation, just like humans, and it is recommended that we should consider the animal effect when we analyze data from safety pharmacology studies.

As stated in section 3, the QT interval is strongly correlated with the HR and RR interval. To evaluate QT intervals that are measured at different HR or RR interval, we conducted a non-clinical study to investigate appropriate QT correction method in cynomolgus monkeys. It was found that traditional correction methods, i.e. Bazett's and Fridericia's methods, are quite biased and could not remove the effect of RR intervals. We proposed individual correction formulas and those methods gave more preferable QTc intervals than traditional methods to evaluate the cardiac safety of drug candidate. Although we need extra baseline/placebo data for each animal to use these individual correction formulas, we can easily obtain such data before we start dosing practically.

Both sections 3 and 4, it was found that there are relatively large individual differences of many parameters from cynomolgus monkeys, and we proposed how to handle the individual differences and between animal variations. Even though, usually, the sample size of safety pharmacology studies is small (n= 4 - 10), and unlike thorough QT/QTc studies in human, a positive control group is not included in a safety pharmacology study and it will not enough to show the sensitivity of the safety pharmacology studies. We need to provide evidence that the study have statistical power to detect the drug effect on QT intervals. And, without information to support appropriate assay sensitivity, it is difficult to know if a compound has some effects on the QT intervals,

or not.

Recently, a new method, called ‘probabilistic analysis method’, was introduced or precisely correcting the QT interval for heart rate. (Holzgrefe *et al*., 2007; Honda *et al.*, 2010). It was reported that probabilistic QT rate-correction method eliminated the confounding effects of heart rate and provided a stable QTc baseline. It also indicated a high degree of sensitivity for the consistent detection of small (5 - 10 msec) changes in the QTc interval. This method is the one of individual correction methods and could show enough assay sensitivity of safety pharmacology studies.

At this time, QT interval is considered as the most important endpoint in the assessment of drug safety (Fermini and Fossa 2003; Finlayson *et al*., 2004), however, there are some clinical and experimental studies showing that QT interval is not an ideal surrogate endpoint. (Hondeghenm *et al.*, 2001; Thomsen *et al*., 2004; Hondeghenm *et al*., 2001). Moreover, there are some drugs that cause QT prolongation without inducing TdP. The value of QT prolongation as a predictor of TdP is further limited, and the association between QT prolongation and TdP seems to be complex. Thus, numerous parameters and ideas have been suggested to be more valuable to quantify risk for proarrhythmia.

Short term variability (STV) of QT interval is one of these parameters applicable in non-clinical and clinical studies for reduced repolarization reserve (Thomsen *et al*., 2004;

Hinterseer *et al*., 2009). Poincaré plots are drawn by plotting each value against the former value of RR and QT intervals (See Figure12 as an example). These plots are often used to visually assess beat to beat variability of repolarization, using some consecutive beats under stable ventricular focus. For a healthy person, it is usually a cloud of points in the shape of an ellipse. (Brennan *et al*., 2001). The mean orthogonal distance from the diagonal to the points of the Poincaré plot is determined and referred to as short term variability (STV):

$$\mathrm{STV} = \frac{1}{n}\sum_{j=1}^{n}\frac{\left|X_{j+1} - X_j\right|}{\sqrt{2}} \quad (j = 1, \ldots, n)$$

where X_j represents *j*-th QT or RR intervals and *n* is the number of consecutive beats.

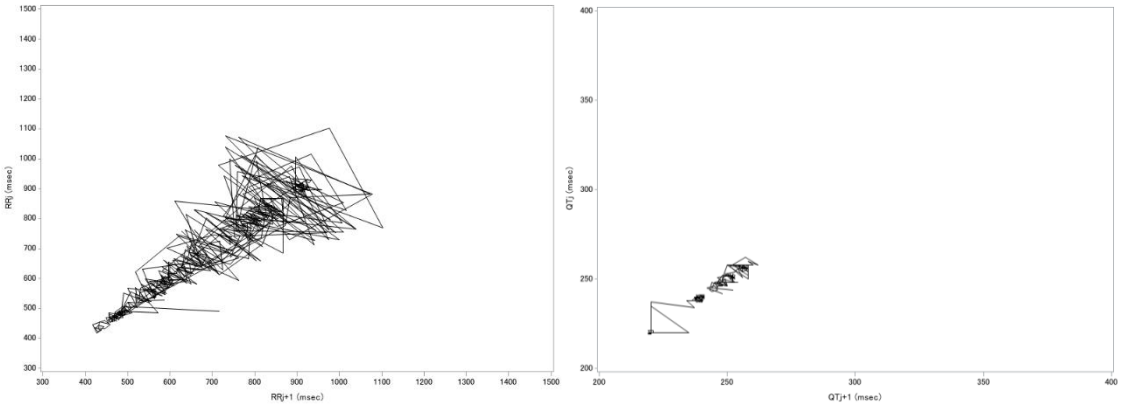

Figure 12: Poincaré plots (Left: RR intervels, Right: QT intervals)

In drug safety assessment, STV for QT or RR intervals is calculated from n consecutive beats (usually 30) at each time point and at each subject/animal. Summary statistics and statistical analyses are performed using STV values. The increased STV, large beat to beat differences in repolarization duration, is considered to due to the susceptibility to proarrhythmia or proarrhythmic potential of drugs.

In safety pharmacology studies, the number of animals is occasionally very limited, so it is reasonable to find the distribution of Z_j first, and then make the required inferences about STV. If we can assume $Z_j=|X_{j+1}- X_j|$, ($j=1, \ldots, n$) are independent and identically distributed, the distribution of STV could easily be computed from the distribution of Z_j. Suppose Z_j $iid \sim gamma$ (α, β), then STV has the gamma distribution with parameters ($n\alpha$, $\beta/n\sqrt{2}$), and similarly the mean and variance of STV are given by $\alpha\beta/\sqrt{2}$ and $\alpha\beta^2/2n$. Gamma distribution belongs to the exponential family; therefore, we can perform statistical analyses based on the generalized linear model (GLM), which is a flexible generalization of ordinary linear regression. GLM covers many situations by allowing for response variables that have arbitrary distributions rather than normal distributions and for an arbitrary link function to vary linearly with the predicted values. Further details about GLM are available in text books such as McCullagh and Nelder (1989), Henrik and Poul (2011).

The beat-to-beat variability of the QT interval has received attention as an independent marker of myocardial vulnerability. The short term QT variability over 30 consecutive intervals predicts d-Sotalol-induced TdP's in dogs. (Thomsen *et al*., 2004) After DXR infusion QT-variability increased suggesting an effect of DXR on the repolarization reserve in humans (Ritsema *et al*., 2009). Further researches about STV in cynomolgus monkeys will be needed as an alternative endpoint to investigate a risk for the development of arrhythmias.

Since the mechanisms of drug-induced QT prolongation and torsade de pointes have not been defined, another endpoint could be extreme QT interval values that focus on and quantify the stochastic behavior of extreme values. Only a basic theoretical framework of an extreme value model is given here, so for more advanced theories, refer to Coles (2001), Kotz and Nadarajah (2000).

Let $M_n = \max\{X_1, \ldots, X_n\}$, where $X_1, \ldots, X_n$, is a sequence of independent random variables, e.g. QT intervals, having a common distribution function F. If there exist sequences of constants $\{a_n>0\}$ and $\{b_n\}$ such that $\Pr\{(M_n-b_n)/a_n \leq z\} \rightarrow G(z)$ as $n \rightarrow \infty$. For a non-degenerate distribution function G, then G is a member of the generalized extreme value (GEV) family of distributions

$$G(z) = \exp\{-[1 + \xi(z-\mu)/\sigma]^{(-1/\xi)}\}, \ -\infty < \mu < \infty, \sigma > 0, \ -\infty < \xi < \infty$$

defined on $\{z: 1 + \xi(z-\mu)/\sigma > 0\}$.

The Fréchet family and Weibull family correspond to the cases $\xi > 0$ and $\xi < 0$ in this theorem. The Gumbel family is considered as the limit of $G(z)$ as $\xi \to 0$. See proofs of the theorems in Leadbetter *et al*. (1983), and the GEV parameterization as in Jenkinson (1955). For a large value of n, the use of the GEV family is an approximation for modeling the distribution of maxima of long sequences. Estimates of extreme quantiles of a block maximum distribution are obtained by inverting $G(z)$:

$$z_p = \begin{cases} \mu - \sigma\,[1-\{-\log(1-p)\}^{-\xi}]/\xi, & \xi \neq 0 \\ \mu - \sigma \log\{-\log(1-p)\}, & \xi = 0 \end{cases}$$

where $G(z_p)=1-p$. z_p is exceeded by the a block maxima with probability p. Since we can obtain a number of ECG data through a telemetry system in cynomolgus monkeys easily, further researches are strongly desired to understand statistical aspects of the extreme QT interval in cynomolgus monkeys.

As in the ICH E-14 guideline, establishing the relationship for experimental drug concentrations to changes in QT/QTc interval may provide additional information to assist interpretation of non-clinical studies assessing cardiac repolarization. Exposure-response analysis is generally recognized that a key component of cardiac safety assessment. To facilitate the evaluation for the exposure-response relationship, pharmacokinetic samples need to be collected at the time of ECG measurements, which

is considered not practical in safety pharmacology studies. However, in safety pharmacology studies, higher doses can be studied and the chance of detecting an effect on the QT intervals can be increased. Exposure-response analysis in safety pharmacology studies can be applied to predict expected effects in clinical studies (Cavero, 2007).

The potential for standardized, integrated, and enhanced non-clinical studies focused on assessing the proarrhythmic potential of a new chemical entity to replace the thorough QT/QTc are currently under discussion and still remains to be established. (Leishman *et al.*, 2012). Enhancing the predictivity of non-clinical studies is likely to involve establishing new strategy. More researches are needed to obtain enough evidences and realize this potential.

Acknowledgments

I would never have been able to finish this doctoral thesis without the help, support, and big encouragement of people around me, to only some of whom it is possible to give particular mention here.

First and foremost, I offer my sincerest gratitude to my supervisors; Dr. Hisashi Noma, Dr. Ikuko Funatogawa, and Professor Koji Kanefuji, for their excellent guidance, support, and great patience. They have provided me with an excellent research atmosphere. I attribute the level of my PhD degree in Statistics to their continuous support and without them this thesis would not have been completed. I wish to especially thank Professor Koji Kanefuji for his support since the days I began working on my PhD research. He helped me come up with the thesis topic and guided me throughout my PhD candidate. His continuous insightful comments and supports has helped me to build my confidence in research.

Besides my advisors, I would like to thank the rest of my thesis committee: Dr. Masahiko Gosho and Professor Yoshinori Kawasaki, for their insightful comments and encouragement, but also for the hard question which incented me to widen my research from various perspectives.

Appreciation also goes out to Professor Satoshi Kuriki, Professor Hiroe Tsubaki, and Professor Yasumasa Baba for all of statistical advices throughout my PhD candidate, and to the office staffs for all helped me along the way. And, thanks also goes out to Dr. Takahiro Sumiya, Dr. Masashi Goto, Professor Yoshimichi Ochi, Dr. Yoshisada Shibata, and Dr. Takao Shohoji who provided me with statistical advice at times of critical need. I

would also like to give many thanks to all friends and colleagues at Clinical Study Support (CSS), Shin Nippon Biomedical Laboratories (SNBL), CPC Clinical Trial Hospital, and Biostatistical Research Association (BRA).

I would like to thank my family. They were always supporting me and encouraging me with their best wishes.

Last but not the least, my greatest thanks to my wife for her personal support and great patience at all times.

Reference

Amador, E. (1974). Normal ranges. *Progress in Clinical Pathology*, 5: 59–83.

Bazett, J.C. (1920). An analysis of time relations of electrocardiograms. *Heart*, 7: 353–367.

Bouvy, J. C., Koopmanschap, M. A., Shah, R. R., Schellekens, H. (2012). The cost-effectiveness of drug regulation: The example of thorough QT/QTc studies. *Clinical Pharmacology & Therapeutics*, 91: 281–288

Box, G. E. P. Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society*, 26: 211–252.

Brennan M, Palaniswami M, Kamen P. (2001). Do existing measures of poincare plot geometry reflect nonlinear features of heart rate variability? *IEEE Transactions on Biomedical Engineering*, 48: 1342–1347.

Cavero I. (2007). Using pharmacokinetic/pharmacodynamics modelling in safety pharmacology to better define safety margins: a regional workshop of the safety pharmacology society. *Expert Opinion on Drug Safety*, 6: 465–471.

Chellman, G. J., Bussiere, J. L., Makori, N., Martin, P. L., Ooshima, Y., Weinbauer, G. F. (2009). Developmental and reproductive toxicology studies in nonhuman primates. *Birth Defects Research Part B: Developmental and Reproductive Toxicology*, 86: 446–462.

Chen, L., Tsong, Y. (2007). Design and analysis for drug abuse potential studies: Issues and strategies for implementing a crossover design. *Therapeutic Innovation & Regulatory Science*, 41: 481–489.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag, London.

David V. G., Larry P. T., Francis W.K. S., Theodore J. B. (2006). Electrocardiogram, hemodynamics, and core body temperatures of the normal freely moving cynomolgus monkey by remote radiotelemetry. *Journal of Pharmacological and Toxicological Methods*, 53: 140–151.

Demacker, P. N., R. W. Schade, R. T. Jansen, A. Van‘t Laar. (1982). Intra-individual variation of serum cholesterol, triglycerides and high density lipoprotein cholesterol in normal humans. *Atherosclerosis*, 45: 259–266.

Démolis, J. L., Kubitza, D., Tennezé, L., Funck-Brentano, C. (2000), Effect of a single oral dose of moxifloxacin (400 mg and 800 mg) on ventricular repolarization in healthy subjects. *Clinical Pharmacology & Therapeutics*, 68: 658–666.

Desai, M., Li, L., Desta, Z., Malik, M., Flockhart, D. (2003). Variability of heart rate correction methods for the QT interval: potential for QTc overestimation with a low dose of haloperidol. *British Journal of Clinical Pharmacology*, 55: 511–517.

Dmitrienko, A. A., Sides, G. D., Winters, K. J., Kovacs, R. J., Rebhun, D.M., Bloom, J. C., Groh, W., Eisenberg, P.R. (2005) Electrocardiogram reference ranges derived from a standardized clinical trial population. *Drug Information Journal*, 39: 395–406.

Ema, M., Arima, A., Fukunishi, K, Matsumoto, M., Hirata, M., Hirose, A., Ihara, T. (2009). Developmental toxicity of dibutyltin dichloride given on three consecutive days during organogenesis in cynomolgus monkeys. *Drug and Chemical Toxicology*, 32: 150–157.

Evans, K., M. F. Laker. (1995). Intra-individual factors affecting lipid, lipoprotein and apolipoprotein measurement: a review. *Annals of Clinical Biochemistry*, 32: 261–280.

Ewart, L., Aylott, M., Deurinck, M., Engwall, M., Gallacher, D. J., Geys, H., Jarvis, P., Ju, H., Leishman, D., Leong, L., McMahon, N., Mead, A., Milliken, P., Suter, W., Teisman, A., Ammel, K. V., Vargas, H. M., Wallis, R., Valentin, J. P. (2014). The concordance between nonclinical and phase I clinical cardiovascular assessment from a cross-company data sharing initiative. *Toxicological Sciences*, 142: 427–435.

Finlayson, F. K., Witchel, H. J., Mcculloch, J., Sharkey, J. (2004). Acquired QT interval prolongation and HERG; implications for drug discovery and development. *European Journal of Pharmacology*, 500: 129–142.

Fermini, B., Fossa, A.A. (2003). The impact of drug-induced QT interval prolongation on drug discovery and development. *Nature Reviews Drug Discovery*, 2: 439–447.

Fridericia, L.S. (1920) Die systelendauer im elektrokardiogramm bei normalen menschen und bei herzkranken. Teil I: Beziehung swischen der pulsfrequenz und der dauer des ventrikelektrokardiogramms bei nermalen menschen in der ruhe. *Acta Medica Scandinavica*. 53: 469–486.

Fukase, H., Fukuzaki, K., Koja, T., Nagata, R., Lukas, S. E. (1994). Effects of morphine, naloxone, buprenorphine, butorphanol, haloperidol and imipramine on morphine withdrawal signs in cynomolgus monkeys. *Psychopharmacology*, 116: 396–400.

Garnett, E. C., Zhu, H., Malik, M., Fossa, A. A., Zhang, J., Badilini, F., Li, J., Darpo, B., Sager, P., Rodriguez, I. (2012) Methodologies to characterize the QT / corrected QT interval in the presences of drug-induced heart rate changes or other autonomic effects. *American Heart Journal*, 163: 912–930.

Gauvin, D. V., Tilley, L. P., Smith Jr., F. W. K., Baird, T. J. (2006). Electrocardiogram, hemodynamics, and core body temperatures of the normal freely moving cynomolgus monkey by remote radiotelemetry. *Journal of Pharmacological and Toxicological*

Methods. 53: 140–51.

Hashimoto, M., Harada, T., Kaga, N., Murano, H., Obata, M. (2002). Accurate evaluation of QT interval in conscious rhesus monkeys (macaca mulatta) by use of holter ECG. *Journal of Electrocardiology*, 35: 333–342.

Hayes, E., Pugsley, M. K., Penz, W. P., Adaikan, G., Walker, M. J. A. (1994). Relationship between QaT and RR intervals in rats, guinea pigs, rabbits, and primates. *Journal of Pharmacological and Toxicological Methods*, 32: 201–207.

Health Canada (2001). Assessment of the QT prolongation potential of non-antiarrhythmic drugs, *Health Canada Therapuetic Products Directorate*.

Henrik M, Poul T. (2011). *Introduction to General and Generalized Linear Models*. Chapman & Hall/CRC, Boca Raton, FL.

Hinterseer M, Beckmann BM, Thomsen MB, Pfefer A, Dalla Poza R, Loeff M, Nets H, Steinbeck G, Vos MA, Kääb S. (2009). Relation of increased short-term variability of QT interval to congenital long-QT syndrome. *American Journal of Cardiology*, 103: 1244–1248.

Holzgrefe H. H., Cavero, I., Gleason C. R, Warner, W. A., Buchanan, L. V., Gill, M. W., Burkett, D. E., Durham, S. K. (2007). Novel probabilistic method for precisely correcting the QT interval for heart rate in telemetered dogs and cynomolgus monkeys. *Journal of Pharmacological and Toxicological Methods*, 55: 159–175.

Honda, M., Komatsu, R., Holzgrefe, H.H., Yamada, Y., Isobe, T., Kimura, K., Itoh, T., Tamaoki, N., Tabo, M. (2010) Application of probabilistic analysis for precisely correcting the QT interval for heart rate in telemetered common marmosets. *Journal of Pharmacological and Toxicological Methods*, 61: 264–270.

Hondeghem, L. M., Dujardin, K.K., Clerck, D. F. (2001) Phase 2 prolongation, in the absence of instability and triangulation, antagonized class III proarrhythmia. *Cardiovascular Research*, 50: 345–353.

Hondeghenm, L. M., Carlsson, L., Duker, G. (2001). Instability and triangulation of the action potential predict serious proarryhythmia, but action potential duration prolongation is antiarrhythmic. *Circulation*, 103: 2004–2013.

Horii, I., Kito, G., Hamada, T., Jukuzono, T., Kobayshi, K., Hashimoto, K. (2002). Development of telemetry system in the common marmoset-cardiovascular effects of astemizole and nicardipine. *Journal of Toxicological Sciences*, 27: 123–130.

ICH Expert Working Group (2000). *Safety Pharmacology Studies for Human Pharmaceuticals S7A*. International Conference on Harmonization.

ICH Expert Working Group (2005). *The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs E14*.

International Conference on Harmonization.

ICH Expert Working Group (2005). *The non-clinical evaluation of the potential for delayed ventricular repolarization (QT interval prolongation) by human pharmaceuticals S7B*. International Conference on Harmonization.

Ihara, T., Funato, M., Chihaya, Y, Honda, T., Nagata, R., Honma, M. (1993). Single and repeated oral dose toxicity studies with talipexole dihydrochloride (B-HT 920 Cl2) in cynomolgus monkeys. *Pharmacometrics*, 46: 401–416.

Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 81: 158–171.

Jones, B., Kenward, G. K., (2003). *Design and Analysis of Cross-Over Trials*. Chapman and Hall/CRC, Boca Raton, FL.

Kaputo, M. J. (1972). The use of probability plots and log-transformations in normal range determinations. *American Journal of Medical Technology*, 38: 1–6.

Kashima, M., Sawada, K., Fukunishi K., Sakata, H. H., Tokado, H., Fukui, Y. (2008). Development of cerebral sulci and gyri in fetuses of cynomolgus monkeys (Macaca Fascicularis). II. Gross observation of the medial surface. *Brain Structure and Function*, 212: 513–520.

Koener, J., Valentin, J. P., Willard, J., Park, E. J., Bi, D., Link, W. T., Fiszman, M., Kozeli, D., Skinner, M., Vargas, H. M., Cantilina, L., Gintant, G., Wisialowski, T., Pettit, S. (2013). Predictivity of non-clinical repolarization assay data for clinical TQT data in the FDA database. ACT Annual Meeting Poster P116. *International Journal of Toxicology*, 32: 63.

Koga, T., Kuwano, K., Kito, G., Kanefuji, K. Evaluation of QT interval using a linear model in individual cynomolgus monkeys. (2007). *Journal of Pharmacological and Toxicological Methods*, 55:248–253.

Koga, T., Nakama, K., Kanefuji, K. (2005). Individual reference intervals of hematological and serum biochemical parameters in cynomolgus monkeys. *International Journal of Toxicology*, 24: 377–385.

Kotz, S., Nadarajah, S. (2000). *Extreme Value Distributions: Theory and Applications*. Imperial College Press, London.

Kurata, M., I. Horii. (2004). Blood coagulation tests in toxicological studies - Review of methods and their significance for drug safety assessment. *Toxicological Sciences*, 29:13–32.

Leadbetter, M. R., Lindgren, G., Rootzen, H. (1983). *Extremes and Related Properties of Random Sequences and Series.* Springer-Verlag, New York.

Leishman, D. J., Beck, T. W., Dybdal, N., Gallacher, D. J., Guth, B. D., Holbrook, M., Roche, B., Wallis, R. M. (2012). Best practice in the conduct of key nonclinical cardiovascular assessments in drug development: current recommendations from the safety pharmacology society. *Journal of Pharmacological and Toxicological Methods*, 65: 93–101.

Li, L., Desai, M., Desta, Z., Flockhart, D. (2004). QT analysis: a complex answer to a 'simple' problem. *Statistics in Medicine*, 23: 2625–2643.

Lipicky, R.J (1993). A viewpoint on drugs that prolong the QTc interval. *American Journal of Cardiology*, 72: 53B–54B.

Liu, Y. W, Suzuki, S., Kashima, M., Tokado, H., Fukuzaki, K., Miyajima, H. (2008). Clinical pathology data from cynomolgus monkeys from China in which diarrhea was observed during quarantine. *Experimental Animals*, 57: 139–143.

Malik, M. (2001). Problems of heart rate correction in assessment of drug-induced QT interval prolongation. *Journal of Cardiovascular Electrophysiology*, 12: 411–420.

Massod, M. F. (1977). Nonparametric percentile estimate of clinical normal ranges. *American Journal of Medical Technology*, 43: 243–252.

McCullagh P, Nelder, J. (1989). *Generalized Linear Models, Second Edition*. Chapman and Hall/CRC, Boca Raton, FL.

Miyazaki, H., Tagawa, M. (2002). Rate-correction technique for QT interval in long-term telemetry ECG recording in beagle dogs. *Experimental Animals*, 51: 465–475.

Molnar, J., Weiss, J., Xhang, F., Rosenthal, J. E. (1996). Evaluation of five QT correction formulas using a software-assisted method of continuous QT measurement from 24-hour halter recordings. *American Journal of Cardiology*, 78: 920–926.

Moritz, A., Y. Fickenscher, K. Meyer, K. Failing, D. J.Weiss. (2004). Canine and feline hematology reference values for the ADIVIA 120 hematology system. *Veterinary Clinical Pathology*, 33: 32–38.

NCCLS. (2000). *How to Define and Determine Reference Intervals in the Clinical Laboratory; Approved Guideline – Second Edition.* Clinical and Laboratory Standards Institute, Wayne, PA.

Qin, T., Shohoji, T., Sumiya, T. (1996). Relationship between adult stature and timing of the pubertal growth spurt. *American Journal of Human Biology*, 8: 417–426.

Raunig, D., DePasquale, M. J., Huang, C. H, Winslow, R., Fossa A. A. (2001). Statistical analysis of QT interval as a function of changes in RR interval in the conscious dog. *Journal of Pharmacological and Toxicological Methods*, 46: 1–11.

Ritsema van Eck, H. J., Broeyer, F. J., Herpen G. V., Burggraaf, J., Kors, J. A. (2009). Short-term QT variability: A marker for reduced repolarization reserve in

anthracyclin therapy. *Computers in Cardiology*, 36: 585−588.

Royston, P. (1991). Constructing time-specific reference ranges. *Statistics in Medicine*, 10: 675–690.

SAS Institute, Inc.(1999). *SAS/STAT user's guide. Version 8.* SAS Institute: Cary, NC.

Shah, A., Hajian, G. (2003). A maximum likelihood approach for estimating the QT correction factor using mixed effects model. *Statistics in Medicine*, 22: 1901–1909.

Shah, R. R. (2002) Drug-induced prolongation of the QT interval: regulatory dilemmas and implications for approval and labelling of a new chemical entity. *Fundamental & Clinical Pharmacology*, 16: 147–156.

Spiegelhalter. D. J., Abrams, K. R., Myles, J. P. (2004). *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. John Wiley & Sons, Ltd. West Sussex.

Sumiya, T., Tashima, T, Nakahara, H, Shohoji, T. (2001). Relationships between biological parameters of Japanese growth of height. *Environmetrics*, 12: 367–382.

Tango, T. (1998). Estimation of age-specific reference ranges. *Statistics in Medicine*, 17: 1231–1243.

Thomsen, M.B., Verduyn, S. C., Stengl, M., Beekman, J. D., de Pater, G., Van Opstal, J. M., Volders, P. G., Vos, M. A. (2004). Increased short-term variability of repolarization predicts d-sotalo-induced torsades de pointes in dogs, *Circulation* 110: 2453–2459.

Tsusaki, H., Yonamine, H., Tamai, A., Shimomoto, M., Kuwano, K., Iwao, H., Nagata, R., Kito, G. (2007). Left ventricular volume and function in cynomolgus monkeys using real-time three-dimensional echocardiography. *Journal of Medical Primatology*, 36: 39–46.

Uno, Y., Matsushita, A., Osada, N., Uehara, S., Kohara, S., Nagata, R., Fukuzaki, K., Utoh, M., Murayama, N., Yamazaki, H. (2010). Genetic variants of CYP3A4 and CYP3A5 in cynomolgus and rhesus macaques. *Drug Metabolism & Disposition*, 38: 209–214.

Uno, Y., Sakuraba, H., Uehara S., Kumano, T., Matuno, K., Nakamura, C., Kito, G., Kamataki, T., Nagata, R. (2009). A null allele impairs function of CYP2C76 gene in cynomolgus monkeys: A possible genetic tool for generation of a better animal model in drug metabolism. *Drug Metabolism & Disposition*, 37: 14–17.

Widjaja, A., Morris, R. J., Levy, J. C., Frayn, K. N., Manley, S. E., Turner, R. C. (1999). Within- and between-subject variation in commonly measured anthropometric and biochemical variables. *Clinical Chemistry*, 45: 561–566.

Williams, E. J. (1949). Experimental designs balanced for the estimation of residual effects of treatments. *Australian Journal of Scientific Research*, 2: 149–168.

Wilt, J.L., Minnema, A.M, Johnson, R.F., Rosenblum, A.M (1993). Torsade de pointes associated with the use of intravenous haloperidol. *Annals of Internal Medicine*, 119: 391–394.

Wright, E. M., Royston, P. (1997). A comparison of statistical methods for age-related reference intervals. *Journal of the Royal Statistical Society: Series A*, 160: 47–69.

Yap, Y.G., Camm, A.J. (2003). Drug induced QT prolongation and torsades de pointes. *Heart*, 89: 1363–1372.

Yoshida, T., Suzuki, K., Cho, F., Honjo, S. (1986). Age-related changes of hematological and serum biochemical aalues in cynomolgus monkeys (Macaca fascicular). [Article in Japanese] *Jikken Dobutsu*, 35: 329–338.