Controversies

August 20, 2015

Edward Shorter: The QT interval and the Mellaril story

Comment by Charles M. Beasley, Jr.

In providing my thoughts on the extent to which the FDA's restriction on dosing of citalopram was or was not warranted, I will comment on three sequential topics. First, did the Zivin et al. study (Zivin, 2013) consider an appropriate outcome variable to exclude an excess of the adverse event of interest with high-dose citalopram compared to low dose citalopram? The adverse event of interest is the ventricular tachyarrhythmia, Torsade de Pointes (TdP). TdP is associated with prolongation of ventricular repolarization caused by blockade of the cardiac rapid delayed rectifier potassium channel I_{kr} (interference with the α-subunit, hERG). Among the several suurogate outcome variables compared between treatments in the Zivin et al. study, the surrogate outcome most specific for TdP would be "ventricular arrhythmia" (i.e., all instances of any of the following: paroxysmal ventricular tachycardia, ventricular fibrillation and flutter, ventricular fibrillation, ventricular flutter, and cardiac arrest).

Second, if the Zivin et al. study did not exclude risk,did the FDA data for citalopram establish a reasonable probability of excess risk of TdP associated with doses of citalopram >40 mg/day? These data would include spontaneous reports in FDA's pharmacovigilance database (FAERS – FDA Adverse Event Reporting System) and two "Thorough QT (TQT) studies".

Third, if the data are sufficient to infer a reasonable likelihood of excess risk with doses of citalopram >40 mg/day, what is the magnitude of that risk? The magnitude of risk would be a rate (e.g., #events / $1x10^x$-person-years).

One important contextual point is that regulatory bodies must be concerned with any excess risk that would occur in the anticipated population that might receive an approved drug product. With an antidepressant, that population will be in the multiple tens of millions of patients. There would be an obvious concern with the risk that could be demonstrated on a population-wide basis. There would also be a concern if there were credible evidence of risk on an extremely infrequent individual basis. That concern would exist even if risk could not be excluded or confirmed in large, population-based studies, due to a variety of methodological matters.

The specificity of the surrogate outcome is the extent to which it does not count false positive cases (events other than TdP). The importance of specificity has been emphasized in the statistical literature. Increasing specificity, even if sensitivity is decreased (failing to count true positive cases), increases the statistical power in demonstrating a difference between treatment groups if such a difference exists (Copeland, 1977; O'Neil, 1988; Quade, 1980). In other words, if a difference between treatment groups exists, it is better to miss counting some true positive cases to minimize counting false positive cases. This is intuitive when considering the following example. Suppose we hypothesize that Drug A causes ischemic strokes due to the acceleration of atherosclerotic plaque evolution in the carotid and vertebral arteries. If we perform a prospective or retrospective study, then we would want only to count cases of ischemic strokes to maximize our chances of finding a difference in the event of interest. Hemorrhagic strokes, embolic strokes from cardiac sources, subarachnoid hemorrhages, etc., would be occurring at random with relationship to the treatment group assignment. Therefore, these events would be observed approximately equally in both treatment groups and could be identified as 'strokes' in medical records. If these 'strokes' that are not ischemic strokes are included in the count of our outcome variable, then this misclassification of random events as the event of interest could obscure a potential treatment difference. The same principle applies when we are interested in TdP caused by drug blockade of the I_{kr} channel. Even if the outcome of interest makes up a substantial proportion of the surrogate (e.g., >60%) increasing specificity of the outcome used is quite important to a study's ability to find a difference if one exists. If the outcome of interest is a small proportion of a surrogate outcome, it is virtually impossible to identify a difference in the outcome between two treatment groups even if a difference exists, regardless of sample size.

Likewise, it is virtually impossible to exclude a difference in the outcome between two treatment groups even if a difference does not exist, regardless of sample size.

TdP will manifest itself across a clinical spectrum. This spectrum includes: 1) brief, self-limited runs identifiable only on ECG; 2) symptomatic but self-limited runs (pre-syncope [symptom of lightheadness and/or dizziness due to cerebral hypoperfusion] or syncope); 3) survived fatal arrhythmia (patient resuscitated); or 4) fatal arrhythmia (sudden cardiac death [SCD][1]). Unfortunately, because a diagnosis of TdP requires an ECG tracing, any attempt to define its background rate of occurrence in the general population due to any cause will undercount cases. Self-limited cases, asymptomatic or symptomatic, will be missed. Some extremely symptomatic cases, fatal and resuscitated, will be missed. Fatal cases without ECG monitoring will be missed. Identification of cases with 100% specificity and 100% sensitivity would require a large population to be continuously monitored electrocardiographically. From a practical perspective, some surrogate outcome is a virtual necessary. Even in research on the longitudinal outcome for patients with genetically confirmed long QT syndrome (genetic LQTS), two clusters of clinical events are considered as outcomes. One is a less specific cluster including cardiac events of any type but primarily syncope, aborted (resuscitated or spontaneously remitted) cardiac arrest (ACA), and SCD. The second is a more specific cluster including only ACA and SCD (Goldberg, 2008). In patients with genetic LQTS, at least those up to 40 years of age with lesser likelihood of suffering from clinical significant coronary artery atherosclerotic disease[2], most instances of ACD and SCD will be due to sustained TdP deteriorating into ventricular fibrillation. Studies of the longitudinal outcome in genetic LQTS are sometimes cut off at age 40, to increase the specificity of the outcome being evaluated (Goldberg, 2008).

As noted above, Zivin et al. used "ventricular arrhythmia" as a surrogate outcome variable. To evaluate the specificity of this outcome, we are interested in the rate of occurrence TdP irrespective of cause, and the extent to which the rate of occurrence of TdP contributes to the rate of "ventricular arrhythmia" in the general population. The proportion of all cases of TdP that are due to drug as opposed to other causes such as genetic LQTs is also of interest.

I am aware of only two efforts to define the rate of occurrence of TdP (Molokhia, 2008; Sarganas, 2014) based on medical records rather than extrapolation from pharmacovigilance

activities. While the primary interest of both these studies was the estimation of the rate of ***drug-induced*** TdP both studies can allow an approximate estimation of the rate of all-cause TdP.

Sarganas et al. (Sarganas, 2014) conducted the only prospective, active surveillance study in a 51-hospital (including clinics) network in Berlin, covering virtually all of the Berlin population, between 2008 and 2011. There were over 250 physician liaisons in the hospitals and clinics with the research. These physicians were contacted every 2-4 weeks by phone, fax, or e-mail to identify potential cases actively. Potential cases were patients with QTc prolongation who had experienced an episode of TdP, ACA, syncope, or severe dizziness. Patients who represented potential cases were contacted, informed consent was obtained, medical records were obtained, and interviews with the patients were conducted to obtain a supplemental history. Possible cases of TdP were identified by the following criteria: 1) a Bazett corrected QT ≥450 ms in males or ≥470 ms in females (complete bundle branch block, implanted pacemaker, or defibrillator were exclusionary); 2) any of the following confirmed on review – ECG confirmed TdP, or resuscitation, or syncope, or severe dizziness. Including syncope and severe dizziness could lead to an overestimation of possible cases among those initially identified by the hospitals. There were 170 potential cases reported by the hospitals. There were 88 potential cases eliminated on an initial review of records. The reason for eliminating these 88 cases might have included lack of the required QT prolongation, lack of the required clinical symptom/signs, or age <18 years. I found the text ambiguous concerning reasons for elimination, and it might be that as few as 40 cases could be confirmed to not have the required QTc prolongation or the required symptoms / signs. Although the authors' focus was to estimate the drug-induced rate of TdP, not a rate irrespective of etiology, other contributory factors such as ischemia, heart failure, structural lesions did not contribute to case elimination. There were an additional 24 possible cases eliminated for a variety of reasons including the absence of an interview (eight fatal and eight refusals to be interviewed). This resulted in 58 "validated cases" (170 reported – (88 not meeting a criterion + 24 other reasons including no interview) = 58). Therefore, this count of 58 "validated cases" might represent only 45% (eliminating only 40) to 70% (eliminating 88) of cases of TdP coming to medical attention in this study. TdP was identified as an arrhythmia in 57% of the "validated" cases, and 73% of the "validated" cases had QTc values >500 ms. In 35 of the 58 (60%) "validated" case drug contribution to the etiology of the event was judged "probable" or "possible" (none judged "certain"). Categories of drug-relatedness included

"certain", "probable", "possible", "unlikely", "unclassified", and "unclassifiable" by standardized descriptive criteria. "Possible" is a rather equivocal category, and the authors did not break down the counts of "probable" and "possible" cases from among the 35. Many of the 35 cases involved multiple drugs, and other risk factors were present as well. The authors estimated the sex and age-adjusted standardized rate for the combined sex population over the age of 18 in Germany for drug-related TdP (nonfatal coming to medical attention) to be 3.3 / million-person-years.

The smallest possible number of cases that lacked prolonged QTc and/or symptoms is 40. If only these 40 are eliminated from the 170 possible cases, the following crude rates can be estimated based on the age-adusted rate of 3.3 / million-person-years for the 35 validated and possibly/probably drug-related cases:

- Total "validated" TdP surviving cases (n=58): 5.5 / million-person-years
- Total possible TdP cases, including fatalities (n=130): 12.3 / million-person-years

This study provides data on the rate of TdP, particularly cases possibly drug-related but does not provide data on the rate of the outcome variable in the Zivin et al. study, "ventricular arrhythmia" or direct data on the extent to which the rate of TdP that would contribute to the rate of "ventricular arrhythmia".

Molokhia et al. (Molokhia, 2008) conducted a retrospective study also with the focus on ***drug-induced*** TdP. The study was carried out in Southwest France between 1991 and 2005 in five hospitals, one private clinic, and one cardiac emergency unit. Suspect cases were those with ICD-10 diagnostic codes for ventricular tachycardia, ventricular fibrillation, and SCD. This suspect case definition was relatively similar, although slightly more restrictive than the outcome of Zivin et al. Zivin et al. included ventricular flutter while Molokhia et al. did not and Zivin et al. included cardiac arrest while Molokhia et al. included SCD. SCD would be slightly more restrictive than cardiac arrest. There were 861 suspect cases identified. Exclusion of cases was more extensive than in the Sarganas et al. study as will be described below. The flow chart and text explaining case exclusion in the Molokhia et al. manuscript is ambiguous. However, it would appear that of the 861 suspect cases, only 143 were confirmed as not having a ventricular tachycardia, and 190 were confirmed as not having prolonged QT on initial data review.

Therefore, there were potentially 528 cases irrespective of any risk factors or comorbid acute cardiac disease.

Deaths without informative data, cases of cardiomyopathy, cardiac failure, myocardial infarction, conduction defects, and ventricular dysplasia were also eliminated. This resulted in 106 cases remaining. These 106 case met the following criteria (required only documentation in medical record, not an actual ECG): 1) polymorphic ventricular tachycardia, or ventricular fibrillation, or syncope and polymorphic ventricular premature contractions; and 2) Fridericia corrected QT in males >440 ms or >450 ms in females[3]; and 3) absence of ECG or enzyme evidence of cardiac ischemia in the last 3 months. Genetic LQTS was present in 50 of these cases. Of the remaining 56 cases from among the 106, where adequate records were available, 40 were identified with drug exposure as likely etiology. Of these 40, only 19 had an ECG available for review showing polymorphic ventricular tachycardia or fibrillation and only 9 of these had an ECG showing prolonged QT.

The authors estimated a crude rate, based on the 40 cases, of 10.9 / million-person-years for drug-induced, symptomatic coming to medical attention, nonfatal cases TdP. Working from this figure, we can estimate the following crude rates:

- Total validated[4] TdP surviving cases (n=106): 28.9 / million-person-years
- Total possible TdP cases, including fatalities (n=528): 143.8 / million-person-years
- Cases meeting screening criteria (n=861): 234.6 / million-person-years

The Sarganas et al. study was prospective, and, therefore, initial screening was likely more specific, without sacrificing sensitivity, for TdP, than was the Molokhia et al. study. While the Molokhia et al. study was retrospective and resulted in higher estimated rates, it is probably the one most relevant to assessing the adequacy of the specificity of the Zivin et al. outcome variable of “ventricular arrhythmia”. The Molokhia et al. and Zivin et al. studies were both retrospective and the “ventricular arrhythmia” outcome in the Zivin et al. study was similar to the initial case acquisition criteria in the Molokhia et al. study. In the Molokhia et al. study, the estimated rate of screening events was approximately 235 / million-person-years.

Therefore, based on the Molokhia et al. study, using medical record diagnosis to screen for TdP we might expect a rate of 235 / millon-person-years. Based on the validated cases in these two studies (Molokhia, 2008; Sarganas,2014), irrespective of cause, we might observe

between 5-30 cases / million-person-years of TdP with symptoms or ECG evidence of TdP, surviving, coming to medical attention, and with conventional ECG confirmation of QTc prolongation. There would likely be between 30-145 cases / million-person-years of possible TdP coming to medical attention including fatalities and cases that would not otherwise be confirmed because of lack of ECG. With continuous cardiac monitoring, additional cases of TdP without symptoms or very mild symptoms (e.g., only mild and infrequent palpitations or lightheadedness) would also be detected and some proportion of possible cases would be confirmed as actual cases of TdP.

Screening based on diagnosis could be expected to yield 235 cases / million-person-years. If only those cases were excluded where the absence of QTc prolongation and/or strong evidence of TdP could be confirmed, TdP might contribute a rate of 30-145 / million-person-years.

What is notable in the results of the Zivin et al. study are the rates of the outcome of "ventricular arrhythmia". These rates are 6381, 4384, 3416, 5680, 4292, and 3361 / million-person-years for citalopram and sertraline across three dose groups for each drug (see Table 2 [1]). Zivin et al. used ICD-9 codes rather than ICD-10 codes, included ventricular flutter and used cardiac arrest rather than SCD. These rates are an order of magnitude greater than the screening rate in the Molokhia et al. study. Multiple factors might explain this difference. Although Molokhia et al. and Sarganas et al. result in somewhat different rates for cases of highly likely TdP, 5-30 / million-person-years, and rates differing by a greater magnitude for cases of possible TdP that could not be excluded, 12-145 / million-person years, based on these two studies possible or highly likely cases of TdP would make up a small proportion of cases identified in the Zivin et al. study, perhaps 2% to 4%. Additionally, few of the TdP cases would be due to citalopram, even in the high-dose citalopram group and high-dose citalopram is a risk factor for TdP. The outcome studied, "ventricular arrhythmia" lacked the specificity for the adverse event of interest, TdP, to provide relevant, convincing information about the presence or absence of risk regardless of the sample size.

U.S. regulators and consultants (Stockbridge, 2012) have estimated that even with drugs that do cause TdP, TdP will occur in only between 1/100,000 – 1/1,000,000 persons exposed to these drugs. The lower incidence, 1/1,000,000, would be consistent with the range of the rate of

drug-induced TdP cases that can be estimated from the Molokhia et al. and Sarganas et al. studies.

Such an incidence, while not representing a major public health matter on a population basis can be a major health matter for those individuals affected.

The Zivin et al. study and its predecessor, the Leonard et al. study (Leonard, 2011), comparable in design and conducted in a five-state Medicaid database, do not demonstrate that citalopram, even at a dose of >40 mg / day, noticeably contributes to the overall rates of "ventricular arrhythmia" or "cardiac mortality" or "noncardiac mortality" or "all-cause mortality" (the latter 3 being additional outcomes in the Zivin et al. study). This is extremely reassuring from a large-scale public health perspective for those extremely important outcomes. Unfortunately and as explained above, I do not believe these studies exclude possible causation of TdP. More needles (TdP) could be hiding in a smaller haystack ("ventricular arrhythmia") with high-dose citalopram than with lower-dose citalopram or some other antidepressants.

The second matter is whether the FDA data for citalopram establish a reasonable probability of excess risk associated with doses of citalopram >40 mg/day. These data include spontaneous reports in its pharmacovigilance database and a TQTstudy. I cannot comment directly on the adequacy of the spontaneous report data in supporting such a reasonable probability because I do not know the details of the data or its analysis. Such spontaneous adverse event data suffer from several problems. These problems include: 1) under-reporting (lack of sensitivity) of an unknown magnitude; 2) incorrect reporting (a clinical set of events labeled by the reporter as "X", when that set of clinical events would actually represent "Y" – lack of specificity); 3) high variability in alternative etiologies and confounders across reports; and 4) unknown actual numbers of patients exposed to the suspect drug. All of these caveats said, there are sophisticated statistical tools that can be brought to bear on these large, international databases of spontaneous adverse event reports that can be helpful in providing some evidence of causation. The proportion of events identified by a specific adverse event term (i.e., "TdP") or a cluster of adverse event terms (i.e., "TdP" or "ventricular tachycardia", or "ventricular fibrillation", or "cardiac arrest ", or "SCD") from among all adverse event reports for a drug can be determined. This proportion can be compared to that proportion for all drugs in the database. This methodology can be more helpful when there are comparator drugs available

with certain similar characteristics relative to the drug of interest. The method can be more useful when the comparator drugs have the same primary mode of action, the same indications, and entered the market at approximately similar times (within several decades). If there is a substantial difference in these proportions, while far from proving causation, supportive evidence exists that must be weighed along with other available evidence. Spontaneous adverse event data cannot be dismissed as worthless under some circumstances. When considering events that might be adding 1 event per 100,000-1,000,000 patient exposures to a similar background incidence, these might be the only data pointing to some excess.

With regard to the TQT study data, the first matter is the extent to which a prolonged QTc on the surface ECG, puts a person at risk for TdP or another malignant ventricular tachydysrhythmia (e.g., non-TdP ventricular tachycardia, ventricular fibrillation). This QTc prolongation is influenced by repolarization but is a measure of the entire action potential duration of multiple tissues summed at the body surface. Not all drugs that lengthen the QTc are thought to cause TdP. Hondeghem (Hondeghem, 2006, 2008, 2011a, 2011b; Shah, 2005) has put forward a multifactorial model to explain the genesis of TdP and ventricular tachycardia/fibrillation that includes four factors. These factors are cardiac wavelength, triangulation, reverse use dependence, instability, and dispersion. The latter three factors are referred to as TRIaD.

Cardiac wavelength refers to the product of (conduction velocity * effective refractory period). Conduction velocity is dependent on the slope of Phase 0 of the action potential (steeper slope = greater velocity). The slope of Phase 0 is dependent on the activity of the ion channel primarily responsible for depolarization in a specific cardiac tissue. The effective refractory period is that time from the beginning of Phase 0 (depolarization) through the point in Phase 3 repolarization where the membrane potential reaches about -60 mV. Therefore, the effective refractory period is much of the length of QTc. Cardiac wavelength is essentially reflected in action potential duration that is reflected in QTc length on the surface ECG.

Triangulation refers to prolongation specifically of Phase 3 (repolarization) of the action potential. Total QTc is determined by the time course of Phase 0 through Phase 3 and can be prolonged by slowed Phase 0 through Phase 2 without a slowed Phase 3. Drugs that block I_{kr} will result in triangulation. Reverse use (or rate) dependence refers to a process where blockade

of a cardiac ion channel and its consequences is increased at lower rates of depolarization (heart rates) and decreased at higher rates of depolarization. Instability refers to fluctuations over time in the action potential duration (QTc). Dispersion refers to differences across different cardiac tissues in action potential duration (QTc dispersion) as well as triangulation, reverse use dependence, and instability.

Hondeghem maintains that it is a combination of TRIaD and prolongation of the cardiac wavelength that results in TdP while TRIaD and shortening of the cardiac wavelength will result in ventricular fibrillation. With TRIaD present, QTc lengthening could be protective of an arrhythmia worse than TdP and could be benign without TRIaD present. Some drugs, then that are associated with TdP do not prolong QTc and some drugs that prolong QTc might not put individuals at risk for TdP. However, many non-cardiac drugs that do prolong QTc block the I_{kr} channel in a reverse use-dependent manner and are therefore risk factors for TdP.

While prolongation of the QTc is associated with some false positive and some false negative results with respect to predicting TdP risk, it is a reasonable predictor of such risk because many drugs that cause such prolongation block I_{kr} in a reverse use-dependent manner. How good, then, is the TQT at excluding or establishing such prolongation? The simple answer to this question is that the TQT study is good at excluding effect on QTc, as that is its primary intent. These studies are good in spite of some complexities. These complexities involve: 1) data acquisition (require resolution at the single millisecond level while the width of the ink line on an analogue paper ECG machine used clinically when I was in training can be 40 ms); 2) measurement (the offset of the T-wave is someplace buried in a curve); 3) beat-to-beat variability (perhaps as much as 25 ms even with best acquisition equipment and measurement methodology [Malik, 2001a]); 4) correction of QT for heart rate inadequacies of conventional formulae (the QT-RR relationship is not the equivalent of a law/formula in physics; it varies across individuals, is subject to hysteresis, and can change within individuals with marked change in autonomic tone); and 5) multiple other factors. A single TQT study might cost the sponsor $2-20M of external spend (internal resources adds more to real cost), depending on the number of subjects and complexity of methods necessitated by a number of factors. For those who believe that these studies are of poor quality and do little to define a QTc effect, I would suggest several references

that underscore the understanding of the complexities and the methods employed in these studies to address these complexities (Darpo, 2010; Garnett, 2012; Salvi, 2010; Stockbridge, 2012).

As Stockbridge et al. (Stockbridge, 2012) have pointed out: "The goal of TQT studies is frequently misunderstood. TQT studies do not quantify the risk of drug-induced TdP. TQT studies aim at identifying those drugs that have no involvement in myocardial repolarization. Such compounds can safely be considered as having no danger of drug-induced TdP." The required primary analysis in the TQT study is a non-inferiority comparison with placebo (null hypothesis is that the drug is associated with a greater increase in QTc than is placebo). Malik (Malik, 2001b) provided evidence that an increase in QTc up to 4-5 ms could be observed during placebo treatment. ICH E14 (Anon 2), the document that outlines the basic requirements for the TQT study[5], set the criteria for rejecting the null hypothesis as the 1-sided 95% confidence interval for the difference between drug and placebo in maximum mean change from baseline in QTc being less than 10 ms. Strictly speaking, all that the TQT study can do is support the absence of an effect on QTc or fail to do so. However, when that 1-sided confidence interval exceeds 10 ms and especially when the mean difference itself exceeds 10 ms, an effect is often inferred. Generally with such a mean difference a test with a conventional α and with a null hypothesis of no difference would reject that null hypothesis. In some instances while the non-inferiority test is "passed" (reject the null hypothesis of a difference), a conventional statistical test with a null hypothesis of no difference will also reject that null hypothesis (Beasley, 2005). A drug with a small but statistically significant QTc prolonging effect can be non-inferior to placebo given the definition of non-inferiority. Just because a drug has some QTc prolonging effect, this is not necessarily interpreted as evidence of increased TdP risk. This population central tendency effect, determined most often in very healthy normal volunteers is extrapolated to predict extremely rare individual patient risk. A maximum mean difference from placebo in the 15+ ms range with the bound of the 1-sided 95% confidence interval extending to >20 ms is thought to carry some excess risk that would be detectable in the population exposed to a drug in wide commercial use (multiple 10s of millions of patients). That detection would be possible ***if*** TdP were to be sensitively and specifically identified in that large treatment group and an equal-sized comparator group, not exposed to the suspect drug with the two groups well matched for additional risk factors. Shah, at the time on staff of the British national regulatory agency,

Medicines Control Agency, published a detailed categorization of TdP risk based on maximum mean increase in QTc as follows (Shah, 2002):

Mean maximum QTc Increase Over Placebo	Likely Potential TdP Risk
≤5 ms	None
6-10 ms	Unlikely
11-15 ms	Possible
16-20 ms	Probable
21-25 ms	Almost definite
≥26 ms	Definite

Shah frankly acknowledged: "The difficulties in interpreting such heterogeneous data on mean changes from baseline (referring to data for several drugs that were available to him at the time, some published) when comparing or evaluating drugs are immediately apparent." Nonetheless he wrote: "Based on these and other data on non-torsadogenic drugs, the likely prognostic significance of the placebo-corrected mean peak effects on QTc interval, computed by the author is shown in Table 1 (the table above)." Shah offered no notion of what incidence or rate of TdP might be associated with drugs that conferred risk.

In spite of Stockbridge and colleagues (Stockbridge, 2012) articulating the statistically pristine interpretation of a TQT study where the results fail to support non-inferiority to placebo, FDA officials do take a pragmatic approach to interpreting the potential clinical relevance of substantial mean changes in QTc associated with drug in a TQT study. Temple, Stockbridge, and Laughren (Temple, 2012) stated: "There is generally little concern at 5 ms (upper bound of the 95% CI <10 ms) and substantial concern above a mean of 20 ms." and "An average QT prolongation between 10 and 20 ms represents some concern[3]." (citation is to ICH E14).

Is there any data to support the extrapolation from some magnitude of mean change in a TQT study to rare but present risk of TdP in patients? Pfizer has provided data that are helpful in this area, and they are derived from psychiatric medications (Harrigan, 2004; Miceli 2010). Two "semi-TQT studies" were conducted. They varied from ICH E14 guidance in that placebo treatment was not used. Because first-generation antipsychotics were being administered over multiple days, patients with schizophrenia, rather than healthy volunteers were studied. Because the length of the study was up to 32 days, administration of placebo to patients with

schizophrenia was considered inappropriate. Nevertheless, these studies likely yielded findings that would closely approximate those in a TQT study fully adherent to ICH E14 guidelines.

In the first study (Harrigan, 2004), ziprasidone 160 mg/day, haloperidol 15 mg/day, thioridazine 300 mg/day, risperidone 6-8 mg/day and 20 mg/day, olanzapine 20 mg/day, and quetiapine 750 mg/day were first studied. Metabolic inhibitors were then added to the treatment to result in higher drug concentrations for study. The patients' off-treatment, baseline data were used to calculate a population-specific correction factor for the same mathematical model as used by Bazett and Fridericia (log-linear). The correction factor was 0.35 ($QTc = QT/ (RR^{0.35})$) very similar to the Friderica correction. The results with drug without metabolic inhibition (metabolic inhibition had little influence) were as follows:

Drug	Mean QTc Increase (ms)	1-sided 95% Confidence Interval (ms)[a]
Olanzapine 20 mg/d	1.7	7.1
Risperidone 6-8 mg/d	3.9	7.5
Risperidone 20 mg/d	3.6	10.2
Quetiapine 750 mg/d	5.7	9.7
Haloperidol 15 mg/d	7.1	12.4
Ziprasidone 160 mg/d	15.9	21.2
Thioridazine 300 mg/d	30.1	35.5

[a] Confidence intervals were taken from the Pfizer Briefing Document for FDA Advisory Committee (Anon 3) because they were presented only in graphical form in the published manuscript (Harrigan, 2004)

The aim of the second study (Miceli 2010) was to evaluate the effect of higher doses, and, therefore, higher concentrations / exposure, of ziprasidone and haloperidol without the influence of metabolic inhibitors. The design of this second study was comparable to that of the first study (Harrigan, 2004) but patients had their doses of assigned treatment titrated upward. ECG data were obtained at steady-state on ziprasidone 40, 80, and 160 mg/day and haloperidol 2.5, 15, and 30 mg/day. The correction factor computed in this study was 0.33, identical to that of Fridericia. The results were as follows:

Drug	Mean QTc Increase (ms)	1-sided 95% Confidence Interval (ms)[a]
Ziprasidone 40 mg/d	4.5	7.1
Ziprasidone 160 mg/d	19.5	23.4
Ziprasidone 320 mg/d	22.5	29.4
Haloperidol 2.5 mg/d	-1.2	1.7
Haloperidol 15 mg/d	6.6	11.7
Haloperidol 30 mg/d	7.2	13.1

The results for ziprasidone 160 mg/day and haloperidol 15 mg/day were quite comparable across the two studies contributing to confidence in the accuracy and generalizability of these population mean effects.

I believe that there is a consensus that thioridazine does convey real risk of TdP that can be detected in large populations. The mean observed increase in QTc was 30.1 ms with an upper confidence interval limit of ~36 ms. If placebo-corrected, this value might well have been slightly higher as placebo is often associated with a slight mean decrease in QTc (Anon 4; Beasley, 2011; Loughren, 2013). This point regarding thioridazine conveying real risk began this discussion. So we have evidence that an increase in QTc up to 5 ms or so can represent random variability (Malik, 2001b) and that an increase above 30 ms is associated with a believable risk of TdP (Harrigan, 2004). The 30 ms threshold for clear risk is founded on a belief that thioridazine stands out from the crowd of all other antipsychotics, other than its active metabolite mesoridazine, with respect to TdP risk.

Interestingly, it is difficult to detect a signal differentiating thioridazine from other antipsychotics using SCD as a surrogate for TdP in a large epidemiological database. Ray et al. (Ray, 2009), using the Tennessee Medicaid database used SCD as the outcome variable with access to medical records for clinical review and determination of SCD. The patient population was restricted to ages 30-74 years. This age restriction could have inadvertently degraded the specificity of SCD for TdP because of the increasing likelihood of ischemia being the primary cause of SCD in that adult age range. The primary purpose of the study was to evaluate first-generation and second-generation antipsychotics as groups compared to no antipsychotic use. However, specific drugs were evaluated, and these were haloperidol, thioridazine, clozapine, olanzapine, quetiapine, and risperidone. Compared to non-users, all drugs were associated with a significant increase in the rate SCD. Furthermore, when compared to propensity score matched non-users all drugs except haloperidol were associated with significant increases in the rate of SCD. Importantly, with the propensity-matched results, the confidence intervals for the ratios of the rates for users of all six specific drugs, compared to rates for non-users, overlapped each other. For the six drugs, rate ratios (drug group compared to the non-user group) were compared within drug for low, medium, and high doses (medium and high for clozapine only).

Thioridazine did demonstrate a significant, positive dose-effect relationship, but this was also the case for risperidone (although not as strong as with thioridazine). For thioridazine, there were 15,715 years of patient exposure. This study had difficulty detecting a clear signal for thioridazine that underscores the difficulties with a surrogate outcome that lacks specificity. Nonetheless, thioridazine most likely is associated with real risk of TdP and a mean increase in QTc in the range of 30 ms or more in healthy controls. While a 30 ms almost certainly predicts risk, that risk is quite rare from a total exposure perspective.

Does an increase in QTc that closely approaches 20 ms, as with ziprasidone, predict some risk of TdP? This 20 ms is at or slightly above the mid-point between a 5 ms increase that can be due to random variability and a 30 ms increase that is a real effect and likely predicts risk of TdP. A prospective, 1-year observational cardiac outcomes study (ZODIAC [Strom, 2011]) comparing ziprasidone (N=9,077; 6,198 patient-years) to olanzapine (N=9,077; 6,902 patient years) could not find a differential risk based on "all cardiovascular mortality" by several definitions (not surprising due to lack of specificity) or "sudden death" by several definitions. For analysis of the outcomes while patients were on assigned treatment, and depending on the definition of / criteria for "sudden death", the number of cases with ziprasidone / olanzapine ranged from a low of 1/3 to a high of 25/26. Specificity of the outcome is not high, and the numbers of observed cases of the outcome were low.

For psychiatric medications, data with sertindole offer more evidence that a mean increase in QTc in the range of 20 ms might predict a risk of TdP that can be detected with appropriate case identification that increases specificity. To my knowledge, a TQT study with sertindole has not been performed. However, according to a Lundbeck FDA Advisory Committee Briefing Document (Anon 4), in Lundbeck studies M93-098 and M93-113, placebo treatment was associated with a -6.2±24.7 ms change in QTc with Fridericia correction and sertindole 20 mg/day was associated with a 23.2±30.2 ms change. A dose of 24 mg/day did not result in a greater increase. A European post-marketing observational study, EPOS (Kasper, 2010) found no more cardiac fatalities in the sertindole-treated patients compared to patients not treated with sertindole. The results of the very large, prospective randomized (sertindole versus risperidone) trial with the primary outcomes being all-cause mortality and cardiac events requiring hospitalization provide additional useful data. The publication (Kasper, 2010) reported

no significant difference in cardiac events requiring hospitalization (sertindole 10, risperidone 6; HR = 1.73 [95% CI: 0.63-4.78]), of which 3 were arrhythmias with sertindole and 1 arrhythmia with risperidone. There were more cardiac deaths with sertindole as assessed by both investigators (non-significant) and by an independent safety review committee: 17 vs. 8 (HR = 2.13 [95% CI: 0.91-4.98]) and 31 vs. 12 (HR = 2.84 [95% CI: 1.45-5.55]). The FDA Review Report for the Advisory Committee (Anon 5) provides additional, more specific data regarding potential cases of TdP. The independent safety review committee had identified cases of death, from among those that had been classified as cardiac, that were "sudden, unexpected death". The definition used here extended out to up to 24 hours following the onset of symptoms rather than within 1 hour. Using all "cardiac, sudden, unexpected deaths" during treatment or within 30 days of end of treatment (as was the case for all analyses) and removing patients assigned to risperidone who had sertindole added to their assigned treatment, patients assigned to sertindole who had risperidone added to their assigned treatment, and patients assigned to either treatment who had thioridazine, mesoridazine, ziprasidone, or pimozide added to their assigned treatment, sertindole had significantly more cases than risperidone: 13 vs. 3 (HR = 5.102 [95% CI: 1.453-17.913])[6]. Some evidence appears to exist to support the belief that an increase in QTc in the range of 20 ms does convey the risk of TdP. Data based on restricting the definition of SCD to death within one hour of onset of symptoms and considering only SCD while taking assigned medication or within serval day of discontinuing medication would have been of greater interest.

What are the QTc data for citalopram? These data were published in an FDA authored response (Temple, 2012) to a criticism of FDA's removal of doses >40 mg/day dose of citalopram from approved labeling based on the Leonard, et al. epidemiological study. Two TQT studies were conducted, one with citalopram and one with escitalopram. The following are the summary mean change results. Although it is not stated, these are presumably maximum mean changes, and they actually represent a difference from placebo rather than simply being within treatment changes from baseline. Moxifloxacin was the required positive control necessary to demonstrate assay sensitivity. The upper bound of the 2-sided 90% confidence interval is equivalent to the 1-sided 95% confidence interval. These were 4-way crossover studies with 119 subjects in the citalopram study and 113 subjects in the escitalopram study (very large sample sizes for most TQT studies, driving down the confidence intervals). The middle doses (i.e., 40 mg, 20 mg) were not studied; rather the values are from pK-pD modeling:

Table 1: Citalopram and Escitalopram: Dose-dependent Change in Corrected QT Interval (QTc)[a]

Citalopram		Escitalopram	
Dose	**Change in QTc (90% Confidence Interval) (ms)**	**Dose**	**Change in QTc (90% Confidence Interval) (ms)**
20 mg	8.5 (6.2, 10.8)	10 mg	4.5 (2.5, 6.4)
40 mg*	12.6 (10.9, 14.3)	20 mg*	6.6 (5.3, 7.9)
60 mg	18.5 (16.0, 21.0)	30 mg	10.7 (8.7, 12.7)
Moxifloxacin 400 mg	13.4 (10.9, 15.9)	Moxifloxacin 400 mg	9.0 (7.3, 10.8)

[a] The table has been extracted from reference 33

Given the results, the sponsors of citalopram/escitalopram would have little incentive to publish the study. Journals might have little interest in publication for a variety of reasons including the fact that this study failed to support its primary hypothesis – non-inferiority. The results of this study are conceptually equivalent to an efficacy study in which a putative antidepressant fails to demonstrate efficacy

My personal belief is that the removal of doses >40 mg/day of citalopram from approved labeling was justified. I believe the risk:benefit balance of a dose >40 mg/day dose is tipped very slightly toward the risk side. The 60 mg/day dose is probably on the threshold of being associated with a mean change in QTc in healthy volunteers that predicts some rare risk of TdP. There might be excess reports of cases suggestive of TdP relative to reports about comparable treatments, assessed by appropriate methods, in regulatory databases. FDA has suggested as much, citing published reports (Temple, 2012). If a dose of citalopram >40 mg/day was more efficacious relative to lower doses or other antidepressants, my opinion would be different. Greater efficacy could be a substantially greater remission rate than with lower doses or other antidepressants in the unselected population of patients with Major Depression or patients not remitting on lower doses or other antidepressants. My opinion would also be different if other SSRIs were not available, or antidepressants with other mechanisms action were not available. I am not aware of data that support greater efficacy for a >40 mg/day dose relative to either lower doses or other antidepressants in an unselected population or greater efficacy in a population selected for lack of response to lower doses or other antidepressants. My arguments are comparable to the justification put forward (Temple, 2012) by Robert Temple, Acting Deputy Director of the Office of Drug Evaluation I (management responsibility over the Divisions of Cardiovascular and Renal Products, Neurology Products, and Psychiatric Products); Norman

Stockbridge, Director of the Division of Cardiovascular and Renal Products; and Thomas Laughrin, at the time Director of the Division of Psychiatric Products.

Even if the risk is rare and might impact only 1 in 100,000 to 1,000,000 persons, I would not want to implicitly advocate safe use through labeling in the individual who might be female, with bradycardia, with hypokalemia and hypomagnesemia due to bulimia, and having occult genetic LQTS (all additive risk factors for TdP). At the very least disclosure of the TQT study results and their potential risk implications was, I believe, mandatory. Lacking evidence of exceptional benefit, removing doses >40 mg/day dose was reasonable. Clinicians can use doses outside labeling. The label restricts the sponsor's activities, not those of the clinician. Obviously, the label can have a chilling effect on clinical practice in our litigious society but we do prescribe off-label in select cases with robust informed consent.

I have endeavored to do two things that I believe are important: 1) point out the critical nature of the specificity of a surrogate outcome being used for a clinical event of interest and provide detail on why this specificity is probably always lacking when the clinical event of interest is TdP; and 2) provide an explanation of the modest support for my belief that a QTc increase in the neighborhood of 20 ms carries some risk of TdP. I would encourage the reader to review the three statistical references (Copeland, 1977; O'Neil, 1988; Quade, 1980) that justify the importance of specificity. Do not take my word for it as this is of critical importance.

I spent the first 16 years of my career in the pharmaceutical industry dealing more with safety matters than efficacy although I worked in development and the last 12 years dealing almost exclusively with safety matters. I worked on methods development and specific, complex questions. Questions regarding serious events almost always involve events of extremely low incidence/rate, sometimes combined with modestly high background rate. Drugs will not make it to the market if the answer to a safety question is simple and scientifically robust when the answer to that question is that the drug causes some clinically significant and severe adverse reaction. By simple and scientifically robust, I mean that event of interest occurs with an incidence/rate with drug that would be statistically significant versus the incidence/rate with placebo or a well-characterized comparator and this finding would be replicated in multiple trials. A certain amount of art and uncertainty will almost always be involved in important safety assessments, particularly for drugs that have made it into clinical use.

If there was no problem with the specificity of the outcome variable in the Zivin et al. study, how robust was the sample size (person-time) for excluding a difference in TdP? Is my statement about art being a major component of many important safety assessment decisions for drugs sufficiently safe to reach the market quite flawed? What would rigorous science require to demonstrate definitively either that a 60 mg/day dose of citalopram is associated with risk of TdP or is not associated with this risk? If we apply the standard applied to efficacy by FDA, ***two*** well-controlled prospective studies with random treatment allocation (citalopram or placebo) that find a statistically significant excess of cases of TdP with citalopram 60 mg/day (demonstrate risk) or as an alternative experiment find statistically significant non-inferiority (demonstrate absence of risk).

Under either hypothesis, risk or no risk, we would need to record ECGs continuously throughout the study participation. This would only need to be a single lead, low fidelity recording, and this is quite possible given current technology. A small, two-electrode patch is attached to the torso with blue-tooth transmission to a Wi-Fi device (smart phone) that periodically transmits to a server for storage and analysis.

What will be the sample size of the studies? We will first make a major, impractical assumption that 100% compliance with treatment and ECG electrode placement as well as 100% study completion will be achieved. Based on the Sarganas et al. and Molokhia, et al. studies, let us consider it a reasonable assumption that we would observe 100 cases of TdP / 1,000,000 person-years in the general population (background rate). This could be an over-estimate by close to an order of magnitude but could also be a 50% underestimate. If we count cases of TdP by ECG without significant symptoms (not unreasonable), then the count of 100/1,000,000 person-years is probably reasonable. For sample size calculation, what would be the excess rate of interest with citalopram in the study with the hypothesis of a difference and what would be the limit on excess rate with citalopram in the study with a hypothesis of non-inferiority? Let us use the same excess proportion for sample-size calculations, 25% excess or 125/1,000,000 person-years. For something life threatening and being studied in a well-controlled experiment, 25% might seem excessive, but let us use that figure as some might suggest 50%.

Again, assuming perfect compliance and 100% completion, running the study for 5 years, and random occurrence across time of TdP events, with 80% power and a conventional α=0.05 to

detect a difference between 500 events in 5,000,000 person-years on placebo and 625 events in 5,000,000 person years on drug, the sample size for each treatment group would be 580,690 (for 90% power, the sample size would be 772,024 per treatment group) (Hintze, 2014). This sample size is based on a primary analysis with Fishers Exact Test comparing proportions that would be appropriate for virtually 100% completion of the 5-year treatment period and the random occurrence of events over time. That would be 2,903,450 years of treatment observation per treatment group for 80% power and 3,860,120 years per group for 90% power.

A more realistic analytical design would account for drop-outs, and we can, therefore, consider survival analysis sample size computation. With a 1-year enrolment period, a 5-year observation period, and the assumption of 40% completion with a constant rate of drop-out in both groups (a very optimistic assumption) PASS-13 (Hintze, 2014) would not compute a sample size for 80% power. With a sample size of 458,544 per treatment group, power would be 21.58%.

Even if we could run the ***two*** study with 100% completion and with 90% power and failed to find a significant difference, the proper interpretation would be limited. All that could be properly said was that the study failed to find a risk of TdP associated with citalopram at 60 mg/day. From a proper interpretive perspective, we would not make the statement that we had demonstrated the absence of risk. The studies were designed to find a risk, not demonstrate its absence. We would need to run ***two*** non-inferiority studies to demonstrate the absence of risk. In a non-inferiority design, the ***observation*** of some excess, interpreted as due to random error, must be tolerated.

For the non-inferiority study, setting the expected background rate and observed rate with citalopram both at 100/1,000,000 person-years, the upper limit on the rate with citalopram to declare non-inferiority at 125/1,000,000 (the observed excess tolerated) with 90% power and α=0.025 (both power and α values are conventional with non-inferiority [1-sided] hypothesis testing) PASS-13 was incapable of computing a sample size even with a 20-year study. The sample size for such a study would be astronomically large.

The greater problem is demonstrating the absence of an effect. However, it would be difficult (impossible from a practical prespective) to demonstrate the presence of an effect in a

rigorous fashion if the effect was rare and the rate of occurrence of the effect was only slightly greater than the background rate such as TdP with a non-atiarrhythmic drug. As I said, most interesting safety matters, critical to some relatively few individuals will defy truly rigorous science. Hopefully, we do the best we can with all the possibly relevant data while attempting to be aggressively objective about the limitations of those data and finding the proper balance and accommodation between the interests of individuals and those of large groups.

End Notes:

1. Definitions of SCD vary slightly across multiple sources but a reasonable integrated definition would be: Death due to cardiac causes, heralded by abrupt loss of consciousness, where death occurs within one hour of the onset of acute symptoms. In retrospective studies, if a death due to cardiac causes was unwitnessed but the patient was found dead within 24 hours of being alive and without symptoms this would be considered a case of SCD. Older definitions vary the allowable time course from onset of symptom to death up to 24 hours in some cases. Presence of known, but stable and asymptomatic heart disease, would not exclude a diagnosis of SCD.

2. The most common cause of SCD in an unselected population is ischemia, as high as 80% (Myerburg, 1997).

3. Reference limits for QTcF are some 10-15 ms less than those based on QTcB using comparable data (Mason, 2007).

4. Not all had actual ECGs available to confirm both TdP and prolonged QT.

5. Although ICH (International Committee on Harmonization, a Committee of major national drug regulatory agencies such as FDA) did not release E14 until 2005 (Anon 2), the scientific branch of the European Agency, the CPMP (Committee for Proprietary Medicinal Products) had issued a forerunner of E14 in 1997 (Anon 1). Health Canada issued a preliminary guidance on evaluation of QT/QTc in 2001 and a combined FDA and Health Canada preliminary guidance was issued in 2002 (Wheeler, 2010). These preliminary guidances were known to Review Division staff within FDA and the first two

TQT studies required by FDA were those by the Division of Urologic and Reproductive Products in 2001 for Vardenafil (Morganroth, 2004) and Tadalafil (Beasley, 2005). Both studies designs were consistent with requirements outlined in ICH E14 not published until 2004-2005.

6. For potential or approved drugs making it through FDA review, at least to the point of an Advisory Committee review, the sponsor's briefing document, and even better, the FDA briefing document if one has been prepared are probably the best sources of data on that drug and the multiple studies performed prior to submission. These documents will contain much greater detail than published academic manuscripts and the document authored by FDA perhaps more objective in interpretation. With some effort in searching, these documents can generally be found on the FDA website.

References:

Anon 1. Points to Consider: The Assessment of the Potential for QT Interval Prolongation by Non-Cardiovascular Medicinal Products [http://www.fda.gov/ohrms/dockets/ac/03/briefing/pubs%5Ccpmp.pdf; Accessed 11-June-2015.

Anon 2. Guidance for Industry: E14 Clinical Evaluation of QT/QTc Interval Prolongation and Proarrhythmic Potential for Non-Antiarrhythmic Drugs. http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm073153.pdf. Accessed 11-June-2015.

Anon 3. http://www.fda.gov/ohrms/dockets/ac/00/backgrd/3619b1a.pdf. Accessed 12-June-2015.

Anon4 http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/PsychopharmacologicDrugsAdvisoryCommittee/UCM161901.pdf. Accessed 12-June-2015.

Anon 5. http://www.fda.gov/ohrms/dockets/ac/09/briefing/2009-4424b1-01-FDA.pdf. Accessed 12-June-2015.

Beasley CM, Mitchell MI, Dmitrienko AA, et al. The combined use of ibutilide as an active control with intensive electrocardiographic sampling and signal averaging as a sensitive method to assess the effects of tadalafil on the human QT interval. JACC. 2005; 46:678-87.

Beasley C, Benson C, Xia J, et al. Systematic decrements in QTc between the first and second day of contiguous daily ECG recordings under controlled conditions. PACE. 2011; 34:1116-27.

Copeland KT, Checkoway H, McMichael AJ, Holbrook RH. Bias due to misclassification in the estimation of relative risk. Am J Epidemiol. 1977; 105:488-95.

Darpo B. The thorough QT/QTc study 4 years after the implementation of the ICH E14 guidance. Br J Pharmacol. 2010; 159:49-57.

Garnett CE, Zhu H, Malik M, et al. Methodologies to characterize the QT / corrected QT interval in the presence of drug-induced heart rate changes or other autonomic effects. Am Heart J. 2012; 163: 912-30.

Goldberg I, Zareba W, Moss AJ. Long QT syndrome. Curr Probl Cardiol. 2008; 33:629-94.

Harrigan EP, Miceli JJ, Anziano R, et al. A randomized evaluation of the effects of six antipsychotic agents on QTc in the absence and presence of metabolic inhibition. J Clin Psychopharmacol. 2004; 24: 62-9.

Hintze, J. PASS 13. Ka ysville, Utah: NCSS LLC, 2014. www.ncss.com.

Hondeghem LM. Thorough QT/QTc not so thorough: removes torsadogenic predictors from the T-wave, incriminates safe drugs, and misses profibrillatory drugs. J Cardiovasc Electrophysiol. 2006; 17:337-40.

Hondeghem LM. Use and abuse of QT and TRIaD in cardiac safety research: importance of study design and conduct. E J Pharmacol. 2008; 584:1-9.

Hondeghem LM. QTc prolongation as a surrogate for drug-induced arrhythmias: fact or fallacy? Acta Cardiol. 2011a; 66: 685-9.

Hondeghem LM, Dujardin K, Hoffman P, et al. Drug-induced QTC prolongation dangerously underestimates proarrhythmic potential: lessons from terfenadine. J Cardiovasc Pharmacol. 2011b; 57:589-97.

Kasper S, Moller HJ, Hale A. The European post-marketing observational sertindole study: an investigation of the safety of antipsychotic drug treatment. Eur Arch Psychiatry Clin Neurosci. 2010; 260:59-68.

Leonard CE, Bilker WB, Newcomb C, et al. Antidepressants and the risk of sudden cardiac death and ventricular arrhythmia. Parmacoepideiol Drug Saf. 2011; 20:903-13.

Loghin C, Haber H, Beasley CM, et al. Effect of atomoxetine on the QT interval in healthy CYP2D6 poor metabolizers. Br J Clin Pharmacol. 2013; 75:549-64.

Malik M, Camm JA, Evaluation of drug-induced QT interval prolongation: implications for drug approval and labeling. Drug Safety. 2001a; 24:323-51.

Malik M. Problems of heart rate correction in assessment of drug-induced QT interval prolongation. J Cardiovasc Electrophysiol. 2001b; 12:411-20.

Mason JW, Ramseth DJ, Chanter DO, et al. Electrocardiographic references ranges derived from 79,743 ambulatory subjects. J Electrocardiol. 2007; 40:228-34.

Miceli JJ, Tensfeldt TG, Shiovitz T, et al. Effects of Oral ziprasidone and oral haloperidol on QTc interval in patients with schizophrenia or schizoaffective disorder. Pharmacotherapy. 2010; 30:127-35.

Molokhia M, Pathak A, Lapeyre-Mestre M, et al. Case ascertainment and estimated incidence of drug-induced long-QT syndrome: study in Southwest France. BJCP. 2008; 66:386-95.

Morganroth J, Ilson BE, Shaddinger BC, et al. Evaluation of vardenafil and sildenafil on cardiac repolarization. Am J Cardiol. 2004; 93:1378-83.

Myerburg RJ, Interian A, Mitrani RM, et al. Frequency of sudden cardiac death and profiles of risk. Am J Cardiol. 1997; 80:10F-9F.

O'Neil R. Biopharmaceutical Statistics for Drug Development. Chapter 13 in Peace KE (Ed.) Biopharmaceutical Statistics for Drug Development. New York: Marcel Dekker, 1988.

Quade D, Lachenbruch PA, Whaley FS, et al. Effects of misclassifications on statistical inferences in epidemiology. Am J Epidemiol. 1980; 111:503-15.

Ray WA, Chung CP, Murray KT, et al. Atypical antipsychotic drugs and the risk of sudden cardiac death. NEJM. 2009; 360:225-35.

Salvi V, Karnad DR, Panicker GK, Kothari S. Update on the evaluation of a new drug for effects on cardiac repolarization in humans: issues in early drug development. Br J Pharmacol. 2010; 159: 34-48.

Sarganas G, Garbe E, Klimpel A, et al. Epidemiology of symptomatic drug-induced long QT syndrome and torsade de pointes in Germany. Europace. 2014; 16:101-8.

Shah RR. Drug-induced prolongation of the QT interval: regulatory dilemmas and implications for approval and labeling of a new chemical entity. Fundamental & Clin Pharmacol. 2002; 16:147-56.

Shah RR, Hondeghem LM. Refining detection of drug-induced proarrhythmia: QT interval and TRIaD. Heart Rhythm. 2005; 2:758-72.

Stockbridge N, Zhang J, Garnett C, Malik M. Practice and challenges of QT studies. J Electrocardiol 2012; 45:582-87.

Strom BL, Eng SM, Faich G, et al. Comparative mortality associated with ziprasidone and olanzapine in real-world use among 18,154 patients with schizophrenia: the ziprasidone observational study of cardiac outcomes (ZODIAC). Am J Psychiatry; 2011; 168:193-201.

Temple R, Laughren T, Stockbridge N. Removal from labeling of 60-mg citalopram dose. Parmacoepidemiol Drug Saf. 2012; 21:784-6.

Wheeler W. Evolution of the thorough QT/QTc study. Regulatory Rapporteur. 2010; 7:21-3.

Zivin K, Pfeiffer PN, Bohnert ASB, et al. Evaluation of the FDA warning against prescribing citalopram at doses exceeding 40 mg. Am J Psychiatry. 2013; 170: 642-50.

Charles M. Beasley, Jr

August 20, 2015