

**Cardiac Safety Biomarkers:
Lessons Learned and Challenges Ahead for QTc,
Troponin, and Other Potential Biomarkers**

Daniel Bloomfield
Executive Director, Cardiovascular Clinical Research
Chair, Cardiac Safety Board
Merck Research Laboratories

Norman Stockbridge
Director, Division of Cardiovascular and Renal Products
U.S. Food and Drug Administration

Draft background paper written to stimulate discussion
for the Institute of Medicine Forum on
Drug Discovery, Development, and Translation
workshop entitled:

"Assessing and Accelerating Development of Biomarkers for Drug Safety"

Washington, D. C., October 24, 2008

IOM Workshop Background

As increased regulatory emphasis is placed on development of biomarkers, the scientific community recognizes the urgency of accelerating preclinical and clinical research on these markers, and developing evidentiary standards for their use. Until recently biomarkers have been evaluated one at the time, and have encountered extensive controversy over their utility. The advent of large scale genomic, proteomic, metabolomic and advanced imaging technologies is changing the landscape. The practical implication of this turn of events is that we have a robust biomedical science enterprise that is at risk of producing a large array of biomarkers for drug development and clinical decision making that lack a substantial evidence base. This has already become an issue in cancer, where a patient/doctor team can purchase multiple prognostic tests, each based on somewhat different arrays of biomarkers. In order to make order out of the impending chaos we will need to develop much more efficient approaches to biomarker evaluation and qualification in animals and humans, and policies to guide the effort will be needed. The scale of effort needed will require much more efficient partnerships among government, academia and the private sector.

The objectives of this IOM workshop are:

- to assess the current state of the art for screening technologies to find off target effects early in development—what have we been able to accomplish, and what remains to be done;
- to develop a prioritized questions list to address remaining obstacles;
- And to discuss how to accelerate the efforts through public and private means.

This is one of three background papers that will assess the community's current ability to predict off-target drug effects, will put forward a list of key questions that remain unanswered, and will consider how development of these markers might be accelerated by public and private means. These papers will serve as the basis for discussion for the main portion of the workshop and will challenge workshop participants and attendees to consider how each individual and each stakeholder might actively contribute to advancing this work.

This background paper on cardiac safety biomarkers was written to address the following objectives:

- Set the stage by reviewing the history of QT as a cardiac safety biomarker including recently published regulatory guidance documents and the development of an infrastructure to promote collaboration between industry, academia, and the FDA with the Cardiac Safety Research Consortium (CSRC);
- Summarize the proceedings of the CSRC Think Tank Meeting and catalog other potential cardiac safety biomarkers;
- Examine a less mature cardiac safety biomarker and the challenges for its development: troponin as a biomarker for cardiotoxicity

Background: QT as a safety biomarker – Issues related to its biology, clinical relevance, measurement

The QTc interval is one of the oldest and best known safety biomarkers used in drug development. The effect of a drug on QTc is used clinically by physicians, is an important part of regulatory decision making and has a major impact on how pharmaceutical companies design and prioritize drug development programs. Compared to newer safety biomarkers under consideration, QTc as a safety biomarker has a number strengths and weaknesses (Table 1).

Table 1.
The Strengths and Weakness of QTc Interval as a Safety Biomarker

	Strengths	Weaknesses
Biology	Knowledge of molecular mechanisms and ion channels Cellular models In vivo models	Weak links between experimental models and clinical events
Clinical experience and relevance	Genetic syndromes (LQT), documented clinical events	Rare clinical events, multifactorial etiologies, unpredictable Insufficient data available to close gap between signal and rare events
Measurable biomarker	Old technology, universally available	Low frequency and low amplitude signal resulting in difficult measurement and poor signal-noise Numerous methods of measurement Measured in static condition
Multisector involvement	Interest from academia, clinical medicine, industry (technology, diagnostics, pharma), regulatory agencies	Lack of harmonization among stakeholders Lack of infrastructure for a coordinated collaborative effort (now addressed by CSRC)

Among the strengths, the measurement of this biomarker is nearly universally available, there is a great deal known about the molecular mechanisms of the ion channels that affect ventricular repolarization, there are a number of well established in vitro and in vivo models, there is substantial clinical experience in patients who have a congenital prolonged QT syndrome, and there is broad multisector interest in advancing our understanding and use of this biomarker. Despite these strengths, the QT interval as a safety biomarker also has a number of critical weaknesses. There is no consensus on the optimal method of acquiring, measuring and analyzing the QT interval due in part to the nature of the signal (low frequency/low amplitude, poor signal-to-noise, intrinsic physiologic variability compounded by a number of important confounding factors). At a more fundamental level, the link between the experimental, models of QT and the occurrence of the rare and unpredictable clinical events is weak. Specifically, the clinical epidemiology has not been done to define specifically the probability of an episode of torsade de pointes or sudden death as a function of QT interval and other clinical outcomes. Moreover, there is neither a clearly articulated paradigm for addressing this issue nor are sufficient data available to close gap between the signal and these events.

Regulatory response: guidance documents, processes and development of tools

Beginning in the early 1990's there appeared a progressively increasing signal for the arrhythmia torsade de pointes in the spontaneously reported post-marketing adverse event data for drugs. This led to the removal of a series of drugs from the market, including terfenadine in 1998, astemizole and grepafloxacin in 1999, and cisapride in 2000. The regulatory response to this public health issue was initially formulated in a "points to consider" document jointly authored by Health Canada and the US FDA. This was followed by a ICH guidance document finalized in 2005. This guidance called for "thorough QT" (TQT) studies of new drugs, to assess their potential for causing torsade. The FDA then put in place an interdisciplinary team to handle review of QT-related protocols and study reports, to ensure a uniform response and the systematic accumulation of experience in this area.

As the regulatory response was being crafted, the FDA made a public appeal for the development of data standards for digital ECG data, with the thought that it was going to be critical to review the ECGs from "thorough QT" studies, since prolongation of QT by only a few percent was considered to be clinically relevant. Such a data standard was developed in 2002 and formally adopted by the HL7 standards organization in early 2003. As the data standard was being finalized, the FDA entered into a Cooperative Research and Development Agreement with Mortara Instruments to develop a web-accessible repository for conforming digital ECG data. This repository came on-line in 2005 and now hosts over 2.5 million digital ECGs collected from over 150 clinical studies.

Effect of regulatory guidance on decision making in drug development

As the ICH S7B and E14 guidance documents were being developed, responses from the pharmaceutical industry were mixed. In general, industry appreciated clarification of the standards for the preclinical and clinical assessment of the effect of a drug on ventricular repolarization. In particular, industry was pleased that E14 created a clear definition of a compound with no QTc risk and made it clear that no further evaluation of QTc would be necessary for these compounds. The concerns expressed by industry related to the E14 guidance were around two issues. First, E14 specified that every systemically available small molecule would require a clinical thorough QTc study (TQT) even if the results of the extensive preclinical studies related to ventricular repolarization outlined in S7B were completely normal. Second, E14 set an extremely high bar for declaring that a compound had no QTc risk – at suprathreshold exposures, a compound had to clearly demonstrate an increase in QTc of less than 5-10 ms in a study that demonstrated assay sensitivity by detecting an increase in QTc of a similar magnitude with a positive control (usually moxifloxacin).

These two issues are primarily focused on a fear that very small signals in QTc would be identified in compounds where there was no theoretical risk, no preclinical evidence and no early clinical evidence of QTc prolongation. The initial lack of understanding of what it means when a compound has a 5-10 ms increase in QTc generated considerable uncertainty in drug development.

- What was the clinical significance of such a small increase in QTc?
- What additional studies would be necessary in later phases of drug development to further clarify the clinical significance of an increase in QTc of this magnitude?
- How will do these additional studies affect the timelines and cost of drug development?
- What is the likelihood that this additional data would be able to offset the perceived risk associated with the small but clearly documented increase in QTc from a TQT study?
- How should a company weigh this potential increase in risk against the potential benefits of a drug?
- How would these issues be described in the label?

The uncertainty around the answers to these questions motivated some pharmaceutical and biotech companies to avoid developing compounds with any potential for this liability. The small magnitude of an increase in QTc (at suprathreshold exposures) defined in E14 as a positive TQT study underscored the limited predictive accuracy of preclinical models. In the process of prioritizing compounds in a portfolio, companies began looking for ways to kill compounds with any potential QTc liability. Any increase in QTc in preclinical studies generated the perception that the compound would face enormous hurdles in drug development. Some companies began to discontinue compounds in

development solely because of an in vitro study demonstrating an interaction with the hERG channel. In addition, as compounds advanced through development, companies feared being penalized for evaluating suprathreshold exposures and attempted to minimize this risk by limiting the maximum doses studied.

With regard to drug development, the ultimate success of these regulatory guidance documents will be realized when there is a shared understanding between pharmaceutical companies and regulatory agencies of the clinical significance of a small increase in QTc interval in the context of the benefits of a new molecular entity. Excessive focus on the biomarker in the absence of true clinical risk has the potential to stifle innovation and may lead to an unfortunate decision to discontinue the development of a drug that has the potential to offer patients substantial benefits that are far greater than the actual risk. One solution to this potential conundrum is to create an environment where regulatory agencies, academics, and industry scientists can collaborate to understand better the link between the safety biomarker (in this case QTc) and the event it is intended to predict (in this case *torsade de pointes*). All parties involved will benefit from better clinical epidemiology and better understanding of how to measure and use a safety biomarker. If successful, this type of collaboration is likely to result in better decision making, decision making which places the risks of a drug in the context its benefits. The enormous potential of this approach is highlighted in the FDA's *Critical Path Initiative*; one example of this type of collaboration is the Cardiac Safety Research Consortium, discussed in more detail below.

Notwithstanding the effect of this guidance on drug development, it is also important to understand how this new regulatory guidance impacts physician decision making. Regulatory agencies and pharmaceutical companies provide information on the assessment of the effect of a drug on the QTc interval to physicians in the context of a product insert or a label. This raises a number of important questions worth considering.

- How do physicians utilize the information in the product label?
- How successful are physicians in measuring the QTc interval when instructed to do so in the label?
- How do physicians make risk/benefit decision for an individual patient?
- Are physicians avoiding potentially beneficial medications because of a fear of a small increase in QTc?
- How do physicians respond to the observation of an increase in QTc interval?
- What is the impact of including new warnings in the labels of drugs that have been used for a long period of time (e.g., methadone)

All of these questions focus our attention on the fundamental limitations to the use of QTc in decision making based on the lack of a strong link between QTc prolongation and the risk of a life threatening event.

CSRC

As the ECG warehouse was coming online, the FDA and Duke Clinical Research Institute initiated a public-private partnership called the Cardiac Safety Research Consortium, to include pharmaceutical companies, clinical research organizations, and academic partners in the effort to leverage the ECG warehouse and associated clinical data for mutual benefit. The Cardiac Safety Research Consortium was formed "To advance scientific knowledge on cardiac safety for new and existing medical products by building a collaborative environment based upon the principles of the FDA's *Critical Path Initiative* as well as other public health priorities." The implementation of the CSRC has faced many challenges related to governance, infrastructure, generating resources (both funds and sufficient effort from individuals involved), intellectual property, antitrust and other legal issues, and getting companies to share data in a collaborative environment. Many of these challenges have been or are in the process of being overcome. Companies have begun to share data and CRSC research teams including scientists from industry, academics, and regulators, have begun to make progress on a number of projects.

Importantly, the CSRC has enhanced communication and education by promoting a dialogue among scientists from academia, the pharmaceutical industry, and regulatory agencies. The CSRC has established common ground and an environment where difficult issues can be discussed outside of formal regulatory channels. These discussions have included how to approach the evaluation of the effects of chemotherapeutic agents and large molecules (antibodies and biologics) on QTc as well as different statistical approaches to the evaluation of the effect of a drug on QTc including concentration-response (PK-QTc) modeling. Recently, a number of pharmaceutical companies have agreed to allow the FDA to share anonymized data from the ECG warehouse to create a meaningful dataset that will enable companies and scientists to enhance the use of old measurements of QTc and develop new measurements of ventricular repolarization. This dataset will also provide the opportunity to understand better the effect of moxifloxacin (the most commonly used positive control in TQT studies) including a better understanding of outliers and non-responders. This has the potential to allow for rich studies in pharmacogenomics that might not be possible in a single company. Taken together, the CSRC in concert with the technological and regulatory advances that have been made over the past few years have the potential to provide incremental improvements the utility of QT as a safety biomarker. But it is not clear, at this point in time, if the CSRC will be able to generate the clinical epidemiology studies and data necessary to provide a more refined link between drug-induced QTc prolongation and the risk of developing *torsade de pointes*.

The question before us now is can we develop an experimental paradigm that will improve our fundamental understanding of the link between other biomarkers

and the other rare events? The CSRC is now poised to make a significant contribution and the next few years will determine if the collaborations within the CSRC generate the datasets necessary to provide meaningful and relevant answers to questions that limit the use of QT as a safety biomarker.

Lessons learned from QT

The development of the regulatory guidance documents on evaluating the potential of drugs to prolong QTc, the technological advances that have enabled the formation of the ECG warehouse, and the healthy dialogue that has taken place between academia, the pharmaceutical industry, and regulatory agencies in the CSRC have taught us a number of important lessons that may be applicable to the development of other safety biomarkers. This result is a fairly complete (but still evolving) system to address a public health issue through regulatory and technical developments. The historical account we give makes this endeavor look like a coordinated response, but that is not so. Several key steps were taken by individuals who recognized what needed to be done next and made sure those steps were taken. The original "points to consider" document had its roots in a document authored by Health Canada's Collette Strnad. The effort to develop a digital ECG data standard involved a team of people from pharmaceutical companies, clinical research organizations, device manufacturers, and academia, but the entire process was initiated by and managed by Scott Getzin of Eli Lilly. The Cardiac Safety Research Consortium largely came into being through the efforts of Christopher Cabell, then at DCRI. Had any of these individuals failed to get involved when and to the extent that they did, the result would most likely have been significant delay and suboptimal.

We conclude from this experience that a more coordinated approach to such problems is important, to ensure that the issues are analyzed appropriately, that regulatory strategies get crafted, that the proper technical infrastructure is developed to collect industry-wide experience, and that the proper public-private partnerships are forged to profit from the aggregate experience.

Path Forward for Other Cardiac Safety Issues

The previous discussion of the recent developments related to QT provides insights into the complexity facing us as we try to develop other cardiac safety biomarkers. Recently (October 6th-7th, 2008 in Bethesda, Maryland), an educational collaboration between the Cardiac Safety Research Consortium (CSRC), the Health and Environmental Sciences Institute (HESI), and the FDA, hosted an open think tank forum integrating pre-clinical and clinical challenges of cardiac safety evaluation of new medical therapeutics. At this forum, faculty experts from academia, industry and FDA will gather to discuss key topics in cardiac safety assessment from pre-clinical and clinical perspectives, with particular focus on the translational "gaps" between these two areas of expertise. There were two plenary presentations that set the stage for the value of the

collaborations promoted by the CSRC and HESI consortia (“Collaboration, Critical Path and Cardiac Safety: The FDA View” -- Doug Throckmorton, “How Can Collaborations in Cardiac Safety Efforts Best Impact the Regulatory Landscape?” – Norman Stockbridge). Organizational updates from HESI and CSRC summarized the challenges and solutions to data sharing process and presented the first proof-of-concept report illustrating the sharing of data from a number of companies in the the CSRC ECG Warehouse. The agenda for this think tank forum was not limited to QT as a safety biomarker and included the exploration of a number of other potential safety biomarkers.

- Cardiotoxicity and troponin: Where do they fit in drug development?
- Pre-clinical and clinical testing for QT proarrhythmia: How do they relate to one another and to risk of life-threatening arrhythmic events?
- Pre-clinical and clinical testing for QT evaluation of non-QT proarrhythmia
- Biologics and large molecules: How do we evaluate proarrhythmia and myotoxicity?
- Risks & benefits developing drugs with safety signals-What are the challenges?
- New horizons of cardiac safety programs: Do we need "thorough" blood pressure, heart rate, platelet and lipid studies?

Some examples of biomarkers that might merit further attention because of their link to cardiac morbidity and mortality include

- Heart rate
- Blood Pressure
- Lipids
- Troponin
- CRP
- BNP
- Ex-vivo platelet aggregation
- Imaging biomarkers (cardiac MR)

It is beyond the scope of this background paper to discuss other potential safety biomarkers in any depth. However, we thought it might be useful to catalog a list of other potential safety biomarkers under consideration and then discuss one biomarker (troponin) to highlight the challenges and path forward.

Troponins and cardiotoxicity

The use of cardiac troponin has been recognized as potential biochemical markers of even subclinical myocardial injury. The use of cardiac troponin for the diagnosis of myocardial infarction is well established and there is a great deal of literature supporting its use in this context. Much less is known, however, about the use of troponins to identify drug induced cardiotoxicity.

Two chemotherapeutic agents – the anthracycline doxorubicin and the humanized monoclonal antibody Trastuzumab -- have been associated with cardiotoxicity. Since the toxicity associated with anthracyclines varies considerably between individuals, the use of cardiac troponin has been suggested as potentially important in treatment planning and monitoring to allow maximum anthracycline dosages without sustaining severe cardiac damage and in the development of preventative strategies to limit cardiomyopathy in later life. The story is complicated by the potential finding that the early LV dysfunction associated with doxorubicin may be reversible in the short term, even though clinical heart failure may appear until much later.

Trastuzumab is used for the treatment of advanced breast carcinoma with overexpression of the HER-2 oncogene and has been shown to improve outcomes by increasing the number of responders and prolonging life in women with surgically resected or metastatic HER-positive breast cancer. Interestingly, preclinical animal studies on mice and monkeys did not reveal cardiotoxicity. However, subsequent clinical trials demonstrated an unexpectedly high incidence of cardiac toxicity. Despite the reversibility of Herceptin-induced cardiac changes in most cases, this toxicity frequently leads to discontinuation of antibody therapy. With this example, the cardiotoxicity has been documented with measures of LV dysfunction; little is known about the use of cardiac troponin in identifying patients at risk for this toxicity.

The paradigm of using cardiac troponin identifies cardiotoxicity after the toxicity has already occurred. A number of important questions remain unanswered with this approach.

- Which cardiac troponin assay should be used?
- When should it be measured and how should it be quantified?
- What is the appropriate threshold to establish that the increase in cardiac troponin will be clinically significant?
- How will that threshold be determined in the context of the potential benefits of the drug?
- What should be done with events that are biochemically detectable but below that threshold and therefore may be clinically insignificant?
- How should investigators manage elevations in troponin in clinical studies?
- Which compounds need to have a cardiac troponin evaluation preclinically?
- Are the preclinical models sufficiently predictive? If not, which compounds warrant a cardiac troponin evaluation in clinical studies?
- How can we define a negative cardiac troponin evaluation? Will a positive control be necessary to determine assay sensitivity? How would a positive control be used?

How can we move forward on these issues? What can we learn from prior experience with QT? The CSRC hopes to develop collaborations between academia, industry, and regulatory agencies that will generate an approach to address these questions. This will require an investment in basic science to better elucidate the molecular mechanisms of cardiac toxicity, preclinical models and clinical data to evaluate the use of biomarkers. A coordinated approach to this problem is important to ensure that scientific issues are addressed appropriately, that regulatory strategies get crafted, that an infrastructure is developed to collect industry-wide experience, and that the proper public-private partnerships are forged to profit from the aggregate experience.

Conclusions

While these advances in the use of safety biomarkers represent a tremendous opportunity, to be truly beneficial, this progress must occur in the context of more wide spread reform in (1) how we evaluate the safety of medical products during development and in the post-marketing environment and (2) how effectively society (patients, physicians, industry, health care systems, lawyers and politicians) generate a shared understanding of the how to approach the risk and benefit of medical products

Potential Questions to Trigger Discussion

- What other safety biomarkers should be added to the catalog?
- What paradigm should be used to validate a safety biomarker – ie how much prognostic data is sufficient to warrant using the biomarker in decision making (regulatory and within industry)?
- How would a new measurement of ventricular repolarization get validated? What dataset would be necessary for it to displace QT?
- What are the key steps necessary to develop a biomarker to evaluate cardiotoxicity? Would you develop adriamycin now (assuming there were no other agents with acceptable safety and efficacy) if you had the preclinical and early clinical data on cardiotoxicity and/or the human data out to 6 months? What data would you need to warrant an investment in a compound that might have cardiotoxicity in a small proportion of patients but is likely to have an overall benefit to the population?
- What steps are necessary to make the most out of a collaboration between industry, academia, and regulatory agencies?
- Is there even potential to develop biomarkers to pick up the drugs that cause an increase in background rates of common events (death, MI, stroke) as with coxibs, rosiglitazone, etc?