The Experience with Voluntary Genomic Data Submissions at the FDA and a Vision for the Future of the Voluntary Data Submission Program

MS Orr[1], F Goodsaid[1], S Amur[1], A Rudman[1] and FW Frueh[1]

Drug developers have been using genomic information in drug development strategies for a number of years, but it was unclear how this information would be reviewed by the Food and Drug Administration (FDA). In order to evaluate the regulatory impact of genomic data in current drug development, a workshop was held in May 2002 to discuss aspects surrounding genomic data submission to the FDA (Figure 1).

In this workshop on pharmacogenetics/pharmacogenomics in drug development and regulatory decision-making, the concept of a "safe harbor" for the submission of pharmacogenomic data was introduced,[1] and shortly thereafter, a draft guidance on pharmacogenomic data submissions was developed. In the following year, a second workshop on "Pharmacogenomics in drug development and regulatory decision-making: The genomic data submission (GDS) proposal" was held to gather public feedback on the FDA Draft Guidance for Industry: pharmacogenomic Data Submissions. This guidance, although clarifying what type of genomic data needs to be submitted to the Agency and when, also introduced the concept of "Voluntary Genomic Data Submissions" (VGDSs), and a new classification system for genomic biomarkers. The FDA released the final "Guidance For Industry: Pharmacogenomic Data Submissions" in March 2005[2] and created a "Genomics at FDA" webportal[3] that provides up-to-date regulatory and background information on genomics.

To date, the FDA has received in the order of 30 VGDSs and held close to 20 meetings with VGDSs sponsors. The first submission was received in mid-2004. Since then, two or three submissions are received each quarter. Although the complexity of the submissions was initially low, more recent submissions contain large and highly complex data sets. In addition to the expanding scientific focus of these submissions, several sponsors have been using this type of informal interaction with the FDA as a stepping stone to present the same or related data in a regulatory context to the FDA later, *e.g.*, in a phase II meeting or in protocol assessments for phase III studies, etc. The philosophy behind this approach is clear: it provides the sponsor of such a study with a good understanding of how FDA will react to the use of this information in the context of regulatory decision-making. For example, questions such as "Has the marker been developed appropriately?", "Were the most critical experimental considerations taken?", "Is the approach to use the marker in a prospective study appropriate?", are discussed in VGDSs meetings.

The principle of voluntary submissions has also found interest in other geographic regions and the first bilateral VGDS project in which both the FDA and European Agency for the Evaluation of Medicinal Products (EMEAs) reviewed the submission and discussed their respective findings in a meeting with the sponsor in 2005. To formalize this process, the two agencies issued in 2006 the "Guiding Principles for Processing Joint FDA/EMEA VGDSs", which describe how bilateral VGDSs are being processed and reviewed.[4]

The areas covered in VGDSs are very broad: VGDSs have been received in the areas of cancer, obesity, depression, hypertension, and Alzheimer's disease, to name a few. Within each of these submissions, the focus has been equally broad, for example, issues such as biomarker qualification, device-related questions, study designs, analysis software, etc., have been discussed (**Figure 2**). Although FDA has had, and

[1]Genomics, Office of Clinical Pharmacology, Office of Translational Sciences, Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, Maryland, USA. Correspondence: MS Orr (michael.orr@fda.hhs.gov)

doi:10.1038/sj.clpt.6100053

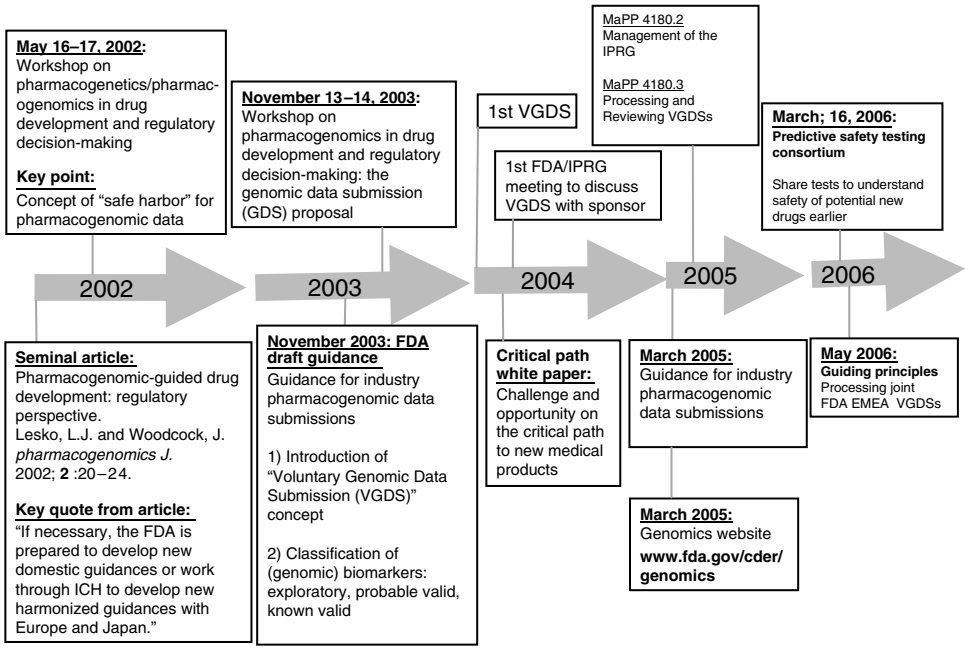

Figure 1 History of Voluntary Genomic Data Submissions (VGDSs) at the FDA.

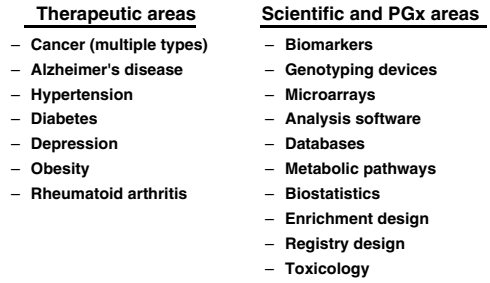

Figure 2 VGDS Types.

continues to have, a great interest in this type of genomic data submissions, the program itself has recently been expanded to VXDS, where "X" stands simply for "exploratory." The idea is to make sure that we do not lose sight of all the other "omics"-type exploratory technologies and that FDA takes an integrated (systems biology) approach to the use of these data in drug development (**Figure 3**). FDA recently received the first of such data sets in which the bridging of these technologies has been the focus.

The development and utilization of the VGDSs at the FDA has added a new catalyst for understanding novel technologies, applications in drug development paradigms, and allows a mutual beneficial scientific forum for drug developers and regulators to discuss the scientific issues associated with cutting-edge technologies and their respective applications. Furthermore, the VGDSs has spurred the FDA to begin developing the appropriate IT infrastructure necessary to store and analyze large genomic data sets. The following section highlights some of the common topics and questions that have been addressed in the VGDSs meetings at this time.

EXAMPLES OF NON-CLINICAL VGDSs

The non-clinical VGDSs received thus far can be grouped into two basic categories:[1] the application of toxicogenomics for the development of screening tools (genomic biomarkers) for target organ-specific toxicities with the goal to select the most promising candidate compound(s) and[2] the leverage of toxicogenomic information along with classical non-clinical information in an attempt to derive more detailed insights into the molecular mechanisms of toxicity.

For example, in the first category, the FDA has received a VGDS from a sponsor that consisted of a TaqMan-based screening system to identify specific subtypes of toxicities in the target organ of interest. Gene expression changes for the sub panels of pre-selected genes were analyzed following treatment of rats with well-characterized reference toxicants that induce specific types of damage. Each study included the classical end points such as clinical chemistry and histopathology results in conjunction with the microarray data. This information was mined for gene expression changes that could identify the particular types of phenotypes/toxicity such as, peroxisome proliferation, necrosis, or apoptosis. It was conveyed by the sponsor that the gene expression information was used in conjunction with preliminary ADME, pharmacological, medicinal chemistry, and toxicology data to help prioritize candidate molecules during early drug development processes.

Questions that were discussed in this context included: "What is the appropriate classification of the genomic biomarkers that they have described in their independent submissions?" Genomic biomarkers as presented in VGDSs to date are exploratory and lack the scientific evidence to move the markers into the "probable valid" or "known valid"

category.[2] Another question was: "Can the FDA comment on a novel set of genomic biomarkers such as selected gene subpanels or signatures that are being used for a particular type of interpretation, *e.g.*, the identification of a PPARα agonist response in a new chemical entity?" Sponsors also have requested feedback from the FDA on the biological interpretation of the gene expression information. In the broader sense, these types of submissions address the issue of how novel e-clinical genomic biomarkers can be qualified for their intended use. The VGDSs program, as well as a series of other activities initiated by the Genomics Group at the FDA (*e.g.*, specific collaborative research efforts and a Predictive Safety Testing Consortium[5]), have led to a proposal for a formal qualification process for such e-clinical biomarkers,[6] which is currently being tested.

Another VGDSs focused on using a toxicogenomic approach for understanding the molecular underpinnings of toxicity for a couple of compounds in their drug development program. One of the many questions posed was: "how does the FDA handle microarray data that had no other supportive findings, *e.g.*, histopathology, associated with it?" The example focused on the presence of increased quantities of peroxisomes that was not detectable via electron microscopy (the gold standard), whereas gene expression changes based on the microarray results provided evidence for increases in peroxisomes. In this case, clearly the absence of confirmatory results for peroxisome proliferation by electron microscopy does not allow the gene expression data to suffice on its own to support the classification for this compound, but the results could be used as part of the evidence to support the presence of increased numbers of peroxisomes.

EXAMPLES OF CLINICAL VGDSs

Different sets of questions have arisen from clinical VGDSs, such as, for example, a study that was targeted at characterizing the efficacy profile of a drug in a patient population using a specific gene variant as a marker. The information submitted was the result of a retrospective analysis of a phase II clinical trial in which a subset of patients containing the gene variant had improved efficacy. These results led to a hypothesis that could be tested in a prospective phase III clinical trial, answering a series of questions, such as: "What is the most appropriate phase III clinical trial design to evaluate the proposed hypothesis?" "Does the data support an enrichment strategy for clinical evaluation?" "Should further clinical evaluation include stratification by genotype?"

Another example of a clinical VGDSs was based on a phase II study in which a gene expression signature was thought to indicate whether a favorable outcome or no benefit could be expected. The sponsor asked if the expression signature could be used for stratification in a phase III trial. The discussion led to the proposal of enrolling all patients (whether they test positive or negative for the signature), and determine whether there were significant differences in outcome in a prospectively defined subgroup analysis of test-positive versus test-negative patients. This study design brought up an interesting question: "Should phase III patients be divided into training and test set populations to generate a new predictive set?" Generally, it was agreed that predictive models can be developed for phase III clinical trials that include a patient population large enough to accommodate a patient distribution that is appropriate to study novel statistical approaches. FDA is currently working on creating several new guidances in this area and a number of different approaches have been proposed and discussed in the recent literature.[7]

It is also noteworthy that during the same time that FDA established the infrastructure and resources to review voluntary submissions, a sharp increase in non-VGDS (albeit of much less complexity) has been observed, indicating that not only the timing for voluntary submissions was critical, but that industry overall has become more comfortable in sharing this type of information with regulatory agencies. Much like the VGDS data in the beginning, the non-voluntary submissions continue to increase in complexity (*e.g.*, from one or a handful of genes that are being studied to whole genome association studies), and it is reasonable to anticipate that in 2–5 years, a significant number of investigational new drugs and new drug applications will contain pharmacogenomic information that is critical to the overall success of the study on hand.

There are also limitations to a voluntary data submission program. Because the submissions are voluntary, they are usually not high priority and they are neither standardized nor are the data bound to be submitted in a specific format. Also, sponsors may use the opportunity to meet "informally" with the regulatory Agency not specifically to discuss the data, but rather to discuss more general aspects of genomics and how the FDA will treat such information. Lastly, FDA may not see all of the data available, as under a voluntary program it is of course possible that not all relevant data is submitted, and asking for more data may not be appropriate in all circumstances.

SUMMARY

The $2\frac{1}{2}$-year-old voluntary genomic data submission program has yielded 30 submissions of increasing complexity. There is a wide variety of therapeutic areas (see **Figure 2**) that were covered and the questions raised included therapeutic area-specific aspects as well as more general biomarker-related and policy questions. Overall, the program has been instrumental to create a review infrastructure for genomic data (hardware, software, intellectual expertise) at the FDA and has led to significant educational and policy-related activities within the Agency (**Figure 1**). The program also has created a novel way to interact with industry on a more informal level, focused on the scientific rather than regulatory interpretation of the results presented. Based on the success of the VGDSs, the voluntary submission program is now expanding so that all other "-omic" (or exploratory)

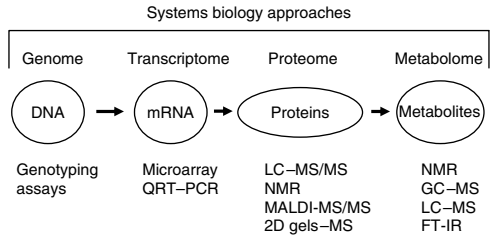

Figure 3 An overview of other "-omic" technologies that could be submitted as a VXDS.

data can be submitted under the new VXDS program (**Figure 3**).

ACKNOWLEDGMENTS

We appreciate the hard work of the FDA Interdisciplinary Pharmacogenomics Review Group (IPRG) members over the last 2 years on Voluntary Genomic Data Submissions to the FDA. The FDA IPRG consists of the following members: Mark I Avigan, Robert Becker, Jennifer Catalano, Maria M Chan, Yvonne Dragan, Bernadette M Dunham, Felix Frueh, James Fuscoe, Steven Galson, Jesse Goodman, Federico Goodsaid, Steve Gutman, Joe Hackett, Indira Hewlett, Shiew Mei Huang, David Jacobson-Kram, Ralph Kodell, Steven, Kozlowski, John K Leighton, Lawrence J Lesko, Marilyn Lightfoote, Robert J Meyer, Shirley Murphy, Robert T O'Neill, Raj K Puri, Allen Rudman, Daniel Schultz, William Slikker, Stephen F Sundlof, Robert Temple, Douglas C Throckmorton, Jawahar Tiwari, Lakshmi R Vishnuvajjala, Sue Jane Wang, and Janet Woodcock.

CONFLICT OF INTEREST

The authors declared no conflict of interest.

1. Lesko, L.J. & Woodcock, J. Pharmacogenomic-guided drug development: regulatory perspective. *Pharmacogenomics J.* **2,** 20–24 (2002).
2. http://www.fda.gov/cder/guidance/6400fnl.pdf.
3. http://www.fda.gov/cder/genomics.
4. http://www.fda.gov/cder/genomics/FDAEMEA.pdf.
5. http://www.fda.gov/oc/initiatives/criticalpath/projectsummary/consortium.hfl.
6. Goodsaid, F. & Frueh, F. Process map proposal for the validation of genomic biomarkers. *Pharmacogenomics* **7,** 773–782 (2006).
7. Freidlin, B. & Simon, R. Adaptive signature design: an adaptive clinical trial design for generating and prospectively testing a gene expression signature for sensitive patients. *Clin. Cancer Res.* **11,** 7872–7878 (2005).