2. The sponsor is using the test results to support scientific arguments pertaining to, for example, the safety, effectiveness, dosing and pharmacology of the drug.

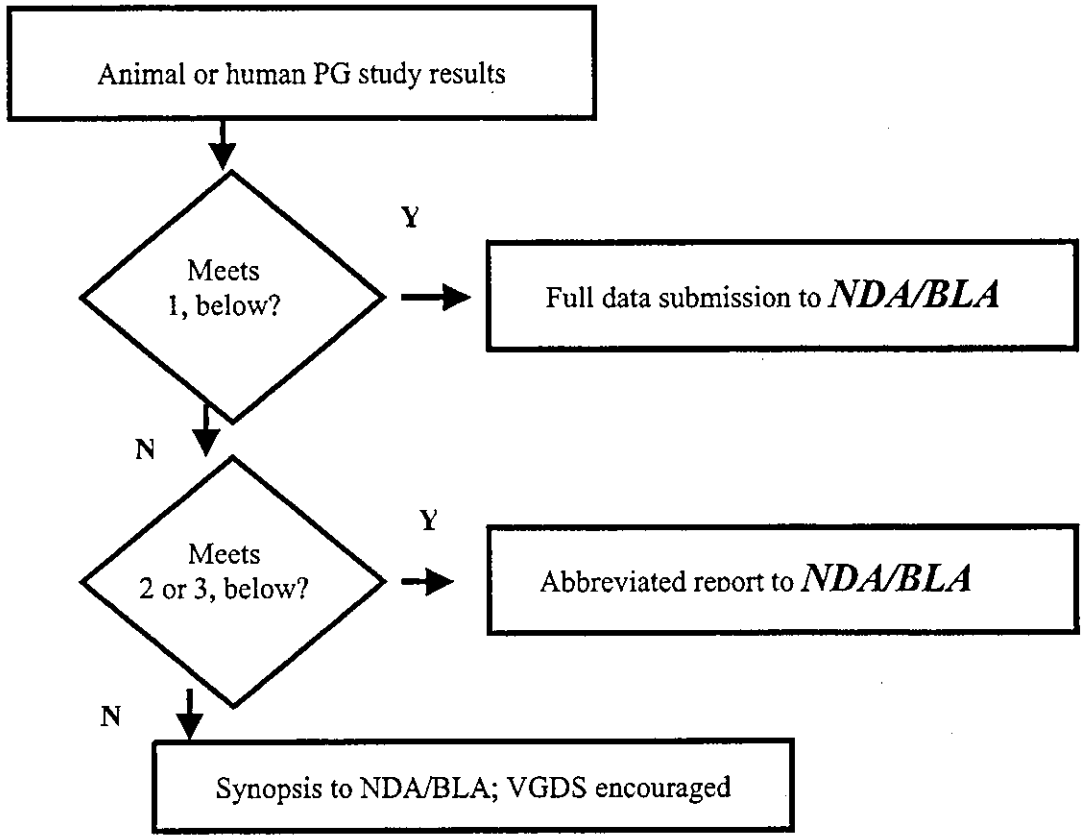
3. The test results constitute a known valid biomarker for physiologic, pathophysiologic, pharmacologic, toxicologic, or clinical states or outcomes in humans, or is a known valid biomarker for a safety outcome in animal studies. If the information on the biomarker (example, human P450 2D6 status) is ***not*** being used for purposes 1 or 2 above, the information can be submitted to the IND as an abbreviated report.

Submission to an IND is NOT needed, but voluntary submission is encouraged (i.e., information does not meet the criteria of § 312.23) if

4. Information is from exploratory studies or is research data, such as from general gene expression analyses in cells/animals/humans, or single-nucleotide polymorphism (SNP) analysis of trial participants.

5. Information consists of results from test systems where the validity of the biomarker is not established.

APPENDIX B: SUBMISSION OF PHARMACOGENOMIC (PG) DATA TO A NEW NDA, BLA, OR SUPPLEMENT



Reports of pharmacogenomic investigations should be submitted to the NDA in the following formats:

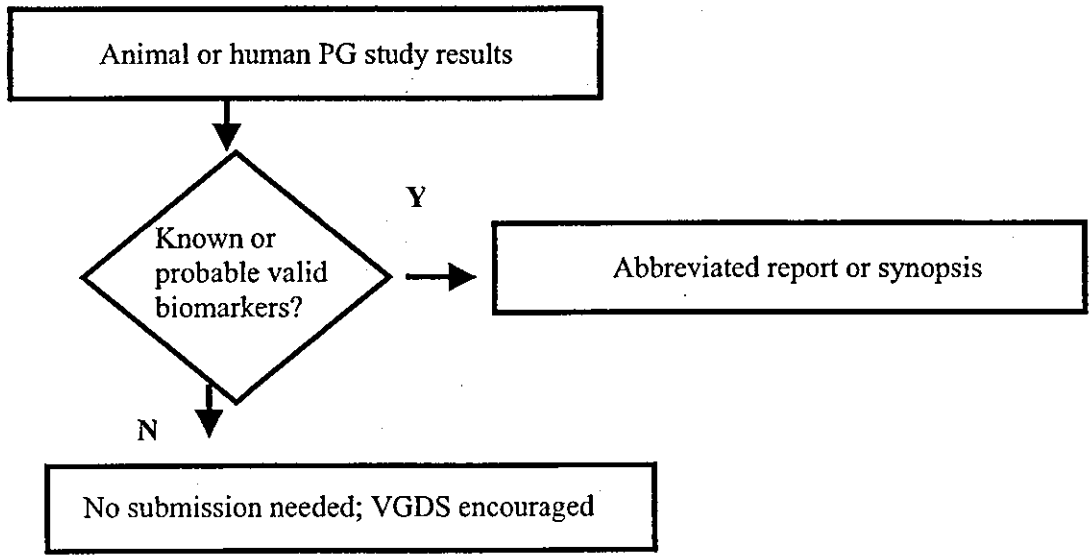
1. Provide reports on pharmacogenomic investigations intended by the sponsor to be used in the drug label or as part of the scientific database being used to support approval as complete submissions (not in the form of an abbreviated report, synopsis, or VGDS), including information about test procedures and complete data, in the relevant sections of the NDA or BLA. If the pharmacogenomic test is already approved by the FDA or is the subject of an application filed with the Agency, information on the test itself can be provided by cross reference.

   The following examples would fit this category.

- Pharmacogenomic test results that are being used to support scientific arguments made by the sponsor about drug dosing, safety, patient selection, or effectiveness
- Pharmacogenomic test results that the sponsor proposes to describe in the drug label
- Pharmacogenomic tests that are essential to achieving the dosing, safety, or effectiveness described in the drug label

2. Submit reports of pharmacogenomic test results that constitute known valid biomarkers for physiologic, pathophysiologic, pharmacologic, toxicologic, or clinical states or outcomes in the relevant species, but that the sponsor is not relying on or mentioning in the label, to the Agency as an abbreviated report (not in the form of a synopsis or VGDS). (If a pharmacogenomic test of this type was conducted as part of a larger overall study, the reporting of the pharmacogenomic test results can be incorporated into the larger study report.)
3. Submit reports of pharmacogenomic tests that represent probable valid biomarkers for physiologic, pathophysiologic, pharmacologic, toxicologic, or clinical states or outcomes in the relevant species to the NDA or BLA as an abbreviated report. (If the pharmacogenomic testing of this type was conducted as part of a larger study, the abbreviated report can be appended to the report of the overall study.)
4. There is no need to submit detailed reports of general exploratory or research information, such as broad gene expression screening, collection of sera or tissue samples, or results of pharmacogenomic tests that are not known or probable valid biomarkers to the NDA or BLA. Because the Agency does not view these studies as germane in determining the safety or effectiveness of a drug, the submission requirements in §§ 314.50 or 601.2 will be satisfied by the submission of a synopsis of the study. However, the Agency encourages the voluntary submission of the data from the study in a VGDS submitted to the NDA or BLA.

APPENDIX C: SUBMISSION OF PHARMACOGENOMIC (PG) DATA TO AN APPROVED NDA, BLA, OR SUPPLEMENT



APPENDIX D: EXAMPLES OF PHARMACOGENOMIC DATA SUBMISSIONS

Some examples of when to provide required pharmacogenomic data submissions versus voluntary (VGDS) genomic data submissions are discussed below.

Metabolizing Enzymes

1. Genotyping CYP2D6 activity in phase 1 human volunteers of various racial and ethnic groups for a new drug where CYP2D6 is the major pathway of metabolism. The PG data may be used to define potential ethnic differences and population-specific dosage regimens.

- CYP2D6 polymorphism is well established as a valid biomarker for drug metabolism enzyme activity
- See section IV.A.2 (complete report) and B.1 (complete report)

2. Genotyping CYP2C19 activity in phase 3 clinical trial patients for a new drug where CYP2C19 is one of the pathways of metabolism. The sponsor may use the information in the labeling.

- CYP2C19 polymorphism is well established as a valid biomarker for drug metabolism enzyme activity.
- See section IV.A.2 (complete report) and B.1 (complete report)

3. Genotyping of CYP3A5 activity in healthy volunteers in a clinical study evaluating the interaction of ketoconazole with a new drug, which is a CYP3A substrate. The data may be used to estimate the relative contribution of the polymorphism to inter-individual variability in AUC.

- CYP3A5 polymorphism is currently not established as a valid biomarker.
- See section IV.A.4 (VGDS encouraged) and B.4 (synopsis; VGDS encouraged)

Transporters

1. Genotyping the MDR1 gene encoding P-gp in phase 1 human volunteers following the completion of a bioavailability study. The data may be used to explore causes of inter-individual variability in AUC.

- These are research data.
- See section IV.A.4 (VGDS encouraged) and B.4 (synopsis; VGDS encouraged).

2. Genotyping MDR1 gene encoding P-gp in a phase 3 trial. The sponsor proposes to use two different treatment regimens based on genotypes.

- Data will be used in clinical decision making (affect dose selection).

- See section IV.A.1 (complete report)

Receptors

1. The sponsor reported that 5-HT1A Ser22 allele is found to be associated with poor response to an SSRI anti-depressant. Individuals with the marker genotype are excluded from the trial to enhance the drug's efficacy profile in a phase 2 proof of efficacy study

- Data will be used in clinical decision making (entry criteria).
- See section IV.A.1. (complete report)

CLINICAL OUTCOMES

Efficacy

1. The sponsor of a monoclonal antibody for treatment of an autoimmune disease has discovered MHC genetic markers predictive of hypersensitivity reactions upon intravenous infusion of the product. The sponsor has also determined that serum concentrations of the antibody 4 weeks after infusion are significantly lower among patients who developed initial infusion reactions. The sponsor genotypes the MHC markers predictive of *infusion* reactions in every patient of a prospective clinical study. It is determined that patients with the genotypes predictive of infusion hypersensitivity (regardless of whether an infusion reaction developed or not) evidence a statistically significantly reduced response to the antibody. The sponsor proposed to highlight the improved efficacy demonstration with genetic stratification in the description of the effects of the drug.

- Data could be used in clinical decision making
- See section IV.A.2 (complete report)
- The sponsor is encouraged to develop a pharmacogenomic diagnostic test (unless it is already available), if it to be reflected in labeling

Safety and Efficacy

1. In a clinical trial, psoriatic lesions are biopsied for gene expression profiling of 160 known disease-associated genes and 140 genes that seemed to correlate with response for the purpose of comparing responders and non-responders to an investigational new drug. Traditional, core clinical measurements are also made to provide evidence of efficacy and safety. The investigation is intended to identify specific gene expression patterns that could possibly be used to correlate with, and predict, efficacy or an adverse event, but at present they do not intend to incorporate the genetic information into labeling.

- These are research data
- See section IV.A.4 (VGDS encouraged).

2. A sponsor filed an IND 3 years ago. During clinical trials, there was lack of efficacy and so the development of the drug was abandoned. Nevertheless the drug had some interesting pharmacological actions that warranted further investigation by the sponsor. The sponsor runs a series of genomic studies in rats and dogs with the drug and discovers a novel pharmacological profile that leads to plans to develop the drug for a different indication.

- These are research data.
- See section IV.A.4 (VGDS encouraged) and B.4 (synopsis; VGDS encouraged)

2.1 Based on the results of the rat and dog pharmacogenomic studies, the sponsor elects to assess a subset of 25 genes in later clinical trials that may be relevant to the safety or efficacy of the compound

- These are supportive data
- See section IV.B.2 (complete report).

Safety

1. Vasculitis is a major drug-related nonclinical safety signal and the basic mechanism of toxicity is unknown. It is normally confirmed by histopathology. A sponsor can use new rat gene chip microarray technology for expression profiling of 8000 known sequenced genes to investigate the mechanism of toxicity and possibly see a pattern of genetic biomarkers in treated rats that is different from controls.

- These are research data
- See section IV.A.4 (VGDS encouraged)

2. A sponsor filed an IND 12 months ago. During the course of subchronic toxicity testing to support longer clinical trial designs, the sponsor finds that rats develop cataracts. This finding represents a safety concern and the sponsor elects to run toxicogenomic studies to define the mechanism of the toxicity. The sponsor discovers that the mechanism is not relevant to humans and uses the data to make their argument about human safety and the absence of cataract risk.

- These are supportive data
- See section IV.A.2 (complete report)

3. A sponsor is investigating a new drug class and seeks to select for clinical development the best of 20 drugs showing some promise in their efficacy screen. No IND has yet been filed. The sponsor elects to assess differences in gene expression profiles to help with prioritization. The data may be generated from animal studies or from cell culture studies. The sponsor feels that the comparative profiles of gene expression alterations between the 20 drugs may help to select the most effective agent with least potential for toxicity. The data are generated to assist with compound selection and are not intended to support the safety of a proposed clinical investigation.

- These are research data
- See section IV.A.4 (VGDS encouraged)

4. A sponsor completes a 2-year carcinogenicity assay in rats and finds that there is an ambiguous tumor signal generated in the kidney, a site that is generally resistant to tumor induction. The sponsor elects to prove that the event was a spontaneous event that was not drug related by dosing the same strain of rats with drug and they succeed in showing that there is no effect of the drug on gene expression in the kidney. A positive control shows a gene expression profile that is very consistent with known pathways of carcinogenesis. The data are used to argue to regulatory authorities that the drug is safe and does not present a tumorigenic risk to humans.

- These are supportive data.
- See section IV.A.2. (complete report)

5. A sponsor conducts global gene expression analyses to assess the relationship between dose and target organ effect. Their drug is a novel acting antipsychotic agent. The sponsor has experience that leads them to suspect that the dose-limiting effect of their drug candidate will be injury to the kidneys - an insidious chronic progressive nephropathy. Using pharmacogenomic analyses, the sponsor finds that reliable and reproducible effects on kidney gene expression occur in both rats and dogs at a dose that is 20-fold lower than the doses in 30-day studies causing a demonstrable histopathology lesion or changes in serum markers for renal toxicity. Insufficient information is currently available to definitively link the more sensitive dose-response changes in gene expression patterns to future changes in renal function or histopathologic lesions.

- These are research data
- See section IV.A.4 (VGDS encouraged)

COMMENTS ON DRAFT GUIDANCE FOR INDUSTRY ON PHARMACOGENOMIC DATA SUBMISSIONS [DOCKET NO. 2003D-0497], FEDERAL REGISTER NOTICE: NOVEMBER 4, 2003 (VOL. 68, NO. 213, 62461-62463)

OVERALL COMMENTS

On November 4, 2003, the Food and Drug Administration (FDA) issued the above referenced Federal Register Notice soliciting public input on draft guidance to industry on Pharmacogenomic Data Submissions. The draft guidance provides recommendations to sponsors holding investigational new drug applications (INDs), new drug applications (NDAs), and biologics license applications (BLAs) on what pharmacogenomic data to submit to the agency during the drug development process, the format of submissions, and how the data will be used in regulatory decision making. The draft guidance is intended to facilitate scientific progress in the area of pharmacogenomics which should enable the FDA to use pharmacogenomic data in regulatory policies and decision making.

The unravelling of the human genome and advances in genetic research are now opening up new horizons in the understanding of the science behind the variability between individuals. GlaxoSmithKline (GSK) is a leader in the conduct of pharmacogenomic research to provide safer and more effective medicines for patients. We applaud FDA for their willingness to partner and work with Industry to develop this guidance as well as for their acknowledgement of the 'state-of- the-art' regarding pharmacogenomics. FDA's intent to issue further guidance on the co-development of pharmacogenomic tests and drugs in the near future is fully supported. It is considered that such guidances are as imperative for Reviewers and the IPRG as they are for Industry if appropriate and consistent use of pharmacogenomic information through provision of a clearly delineated, predictable, process is to be ensured.

The ongoing dialogue between Industry and FDA and activities such as the recent joint workshop to discuss VGDS are welcomed and supported. It is hoped too that FDA will continue to liaise globally for a harmonized approach that is supported by all major regulatory agencies given the potential global regulatory impact of pharmacogenomics on drug development.

GSK believes that the guidance, with suggested modifications, provides a reasonable framework to facilitate scientific understanding and progress in the field of pharmacogenomics through the free exchange of information. Additionally, the guidance is beneficial in promoting the use of certain pharmacogenomic data in informing regulatory decisions for the improved use of medicines.

The provision of a process whereby companies may share exploratory pharmacogenomic data with FDA whilst having the mutual goal of advancing the state of scientific knowledge without impeding the progress and generation of such data, or the availability of new and needed treatments for patients, is welcomed and fully supported. In addition, GSK considers it critical that there is an ongoing information exchange between FDA and Industry to share the educational benefits and insights gained through the VGDS initiative.

It is also believed, however, that there are significant opportunities to further the utility of the guidance, particularly with regard to critical definitions (and related implications) such as biomarker definitions as well as the details pertaining to when and how data will, and will not, be used for regulatory decision making.

The areas where GSK advocates revisions to the guidance are summarized in this document.

GENERAL COMMENTS

?? We suggest that the guidance should be intended for both Industry and FDA (IPRG and Reviewers). **[cover page]**

?? It is assumed that, once issued, the guidance would be applied only to ongoing investigations and new marketing applications. **[lines 139, 278 and 340]**

?? It is unclear why proteomics is excluded from the guidance. It would be helpful for the science to have clarity on handling these exploratory biomarker data, and indeed other 'omics' such as metabolomics, and thus permit sponsors to submit such data for review and discussion under VGDS. **[line 31]**

DEFINITIONS

?? The definitions for pharmacogenetics and pharmacogenomics appear to have unnecessary and confusing overlap i.e. *"interindividual variation in DNA sequence related to [pharmacokinetics] or [pharmacodynamics]"* defined in pharmacogenetics *and "interindividual variations in whole genome or candidate gene single-nucleotide polymorphism (SNP) maps, haplotype markers....."* defined in pharmacogenomics, whereas the latter are actually approaches to conducting the former.

It is suggested that both of the definitions noted above are applied to pharmacogenetics (i.e. related to DNA) and that pharmacogenomics relates specifically to the analysis of gene expression and its products. **[lines 26-33]**

?? The draft guidance makes reference to a "biomarker", "valid biomarker", "known valid biomarker" and a "probable valid biomarker". Which of the three categories the sponsor assigns determines the level of reporting details. Whilst it is

understandable that data used for regulatory decision making should require mandatory submission, the value of implementing a "validity gradient" is questionable and will likely result in confusion and inconsistent application by various sponsors. GSK suggests that the primary point of emphasis for the submission decision algorithm be whether the data are used for regulatory decision making and offers the following suggestion for such a definition:

Pharmacogenomic data used for "regulatory decision making" are data sufficiently established to make assessments regarding the safety and efficacy of the drug (i.e. predictive value) that guide:

- *the sponsor's decisions regarding the design or selection of non-clinical or clinical research studies, or*
- *a regulator's determination of the acceptability of proposed biomedical research, or*
- *a regulator's determination of approvability of a marketing application, or change in the recommended conditions of marketed product use (e.g. labeling)*

Pharmacogenomic data that are used for all other purposes can be considered as not used for Regulatory Decision Making.

GSK would advocate that pharmacogenomic data for biomarkers that are used for regulatory decision making (as defined above) should require mandatory submission to FDA. Studies of all other biomarkers would be encouraged under VGDS for INDs, NDAs, and BLAs.

If FDA determines that the proposal for categories of biomarkers should be retained (described above as a "validity gradient"), we feel that the Agency will need to address the practical considerations of ensuring that all parties share a common understanding and provide clear guidance as to how all may consistently determine the appropriate category for regulatory reporting purposes.

As noted on line 126, the distinction of a biomarker will evolve over time. GSK suggests that if a category, rather than the action taken, is the focus, the only means by which all sponsors can share a common understanding of the regulatory implications of a given biomarker is if FDA were to maintain, and make available publicly, a list of what the Agency considers to be "known valid biomarkers". If this approach is taken, the Agency should replace the term "known valid biomarker" with "approved (or accepted) biomarker" as this would be a more accurate reflection of the relevance to the regulatory process. Such a list should specify the conditions under which the biomarker has been judged "approved". FDA should also include in the category of "approved" those biomarkers that individual sponsors have established with the Agency as

sufficiently clinically significant to be used for regulatory decision making. Submission of all other biomarker data (i.e. that are not used for regulatory decision making) should be encouraged under VGDS.

Even though GSK has proposed the alternative above for an FDA-maintained list, it is recognised that FDA may not wish to pursue this course of action. Therefore, GSK urges the Agency to base its guidance for regulatory reporting on those biomarkers used for Regulatory Decision Making (i.e. the definition proposed previously) rather than a subjective decision about the acceptability of the biomarker to FDA and / or the scientific community at large.

?? Described below are some of the specific issues that will need to be addressed if FDA determines that it wishes to retain the proposed categories for biomarkers:

- There is a need to better define the terms "valid biomarker", "known valid biomarker" and "probable valid biomarker" to aid transparency and predictability for sponsors.

- In addition, the purpose of the "biomarker" (e.g. diagnostic, predictive, prognostic) should also be specified to ensure common understanding with regard to the utility of the information.

- It would be particularly helpful to Industry if representative real examples from FDA were included for what are considered to be known valid and probable valid biomarkers (including for the drug metabolizing cytochrome P450 enzymes), both for the IND phase as well as for the unapproved or approved NDA / BLA phase of the regulatory review of such data. **[lines 121-141]**

- It would be beneficial to Industry if FDA elaborated on what constitutes an *"established scientific framework or body of evidence"* for a valid biomarker with specific examples. Also, It is suggested that the definition of a "probable" biomarker as referenced in the guidance to *"data being generated within a single company"* or *"without independent replication of data"* be further clarified with regard to multi-center / investigator, multi-studies for a given biomarker evaluation and validation assessment. **[lines 130, 138-139 and 607-608]**

- FDA also references a probable valid biomarker as having *".....a significant association between a pharmacogenomic test result and clinical outcomes....."*. It is noted that whilst a significant association may be evident between test results and drug responsiveness, the association with a clinical outcome may be a significant hurdle to clear. FDA is requested to provide additional clarification and examples and also to consider associations with surrogate endpoints in addition to clinical endpoints. **[line 140]**

- FDA is requested to define in more detail the process for the transition from "exploratory" pharmacogenomic data to "probable" valid biomarker and

ultimately to "known" valid biomarker, together with what FDA sees as the consequences and requirements for the sponsor regarding such transitions. **[lines 128-145]**

USE OF PHARMACOGENOMIC DATA IN REGULATORY DECISION MAKING

?? FDA is requested to clarify what is meant by "use in decision making" at both the IND and the NDA / BLA stages, and to provide additional examples of when the data will, and will not, be used in such a manner. GSK would advocate for defining and distinguishing between "Decision Making" and "Regulatory Decision Making". **[lines 62-63, 107-109 and 115-119]**

For example, "observational" (non regulatory decision making) pharmacogenomic data from a given study with a pharmacological agent that are not used to affect or support subsequent study designs / patient selection / stratification / regulatory decisions for that drug, or that are used to make decisions regarding the drug development program of another pharmacological agent, would be eligible for submission under VGDS. However, such data being used "directionally" (regulatory decision making) for a subsequent clinical trial design such as subject selection / screening, or used to support regulatory decision making for that drug, would need to be submitted in full (i.e. not eligible under VGDS). Similarly, in a non-clinical setting, pharmacogenomic data generated and used as a screening methodology for potential pharmacologic or toxicologic activity to better select drug development candidates would be eligible for submission under VGDS.

?? The draft guidance provides a description of those circumstance that would constitute mandatory reporting to FDA of pharmacogenomic data and the expected level of detail to be submitted that is commensurate with the FDA biomarker category and the product registration status **[lines 284-293 and 343-375]**. It is suggested that the descriptions would be more helpful if specific (hypothetical) examples are included in order to illustrate the intent of the descriptions pertinent to non-clinical and clinical data submissions, dosing, efficacy, and safety.

?? We recommend that the guidance reflect that submission of full data sets generated with the microarray technology or SNP association study data is not expected if only an evaluation of a subset of genes is used for regulatory decision making. For example, when research is focused as the result of previous validation experiments, GSK would propose that it is more informative and appropriate that submission of data related only to the subset of genes of interest should be required.

?? The guidance does not address how the Review Divisions will respond to generated data (particularly exploratory data) and under what circumstances. It is requested that FDA outlines the mechanism for the Review Divisions to obtain appropriate (and timely) counsel and input from the IPRG and details safeguards to ensure that Review Divisions do not develop conflicting independent policy decisions. **[lines 498-509]**

?? What is the envisaged process, if any, in the event that there is disagreement between the sponsor and a Review Division regarding what is considered appropriate usage of the pharmacogenomic data? It is suggested that a process for resolving such differences, including the role of IPRG, be outlined in the guidance to facilitate consistency within Industry and across Review Divisions. **[lines 498-509]**

?? The guidance states that where a sponsor develops a drug for a selected population (safety considerations), co-development of an FDA-sanctioned IVD that is available when the drug is marketed, is required; however, where a sponsor is appropriately developing a drug for all-comers whilst also pursuing PG markers for toxicity, then the test could be available as an approved IVD or service.

It is requested that additional clarification be provided for the options highlighted in the draft guidance, recognizing that this may be more appropriately addressed in the 'co-development' guidance that FDA notes is to be available in the near future. **[lines 536-540 and 542-549]**

?? FDA has provided useful guidance regarding GLP data; however, GSK requests that this be expanded. For example, what guidance can FDA provide regarding the desired format / content for such data and what is required for validation? What is FDA's guidance for exploratory data that are generated within GLP non-clinical studies? Also, what is FDA's view regarding pharmacogenomic clinical data that are generated in research labs since these may not possess the same degree of rigorous sample handling / tracking and validated assay methodologies found in *clinical* laboratories. **[lines 393-405]**

?? It would be helpful if FDA provided more precise detail with regard to the scope and format for a full report vs. an abbreviated report vs. a synopsis, together with the conditions under which each is required. **[lines 111-113]**

?? The guidance does not address when FDA would advocate the generation of pharmacogenomic data. It would be helpful if FDA would describe scenarios and the process for incorporating utilization of "approved " biomarkers ("known valid biomarkers") for regulatory decision making into appropriate guidances, as for example, with regard to drug metabolism and CYP2D6. **[lines 106-119]**

VOLUNTARY PHARMACOGENOMIC DATA SUBMISSIONS

?? It is assumed that FDA intends for the sponsor to decide if available pharmacogenomic data fall within the scope of VGDS. What mechanism will be available for consulting with appropriate individuals at FDA regarding the respective sponsor and Agency views regarding such decisions? **[lines 223-242]**

?? There is a need for additional clarity for what FDA will and will not do with the data in VGDS and any assessments made. What is the contact with the Review Divisions and the intent for access to, and use of, the data by the Review Divisions? **[lines 489-509]**

?? Whilst FDA's flexibility regarding the format of data to be submitted under VGDS is welcomed, additional details specifying how much data should be included and what context is required would also be helpful. Also, what is the expectation of FDA for sponsors to meet with the IPRG regarding submitted data under VGDS? **[lines 410-434 and 498-502]**

?? Additional clarification is requested regarding what the 'triggers' could be for *"FDA becoming aware of the significance of a particular PG test after evaluating results across sponsors"* together with what the communication process is for *"notifying sponsors about this determination"*. Combining different sets of VGDS data from different companies runs the risk of erroneous conclusions in addition to conferring significance to a dataset that was deemed initially to be of an exploratory nature – what does FDA envisage as a safeguard in this respect? **[lines 505-507]**

?? FDA is requested to provide additional details regarding the IPRG functioning - who sits on IPRG and will external members to FDA be eligible in a manner similar to the Advisory Committee concept (e.g. NIH, academia). GSK would encourage participation that provides 'state-of-the-art' input and counsel. **[lines 240-242]**

?? Based on the premise that confidentiality of the VGDS data needs to be maintained, how will this be achieved within the IPRG and how will potential conflicts of interest be managed? **[lines 236-242]**

?? GSK considers that information sharing is a critical component and incentive for industry with regard to the VGDS initiative. FDA is requested to outline the opportunities and process for this aspect regarding both an individual sponsor's data with IPRG as well as for 'cross-sponsor' data where a pattern of association might be identified by the IPRG. In the latter situation involving multiple sponsors, GSK supports discussion that is inclusive of all sponsors with the appropriate maintenance of sponsor proprietary information.

NIH Comments on FDA's Draft Guidance for Industry Pharmacogenomic Data Submission Docket No. 2003D-0497

We are writing in response to the request for comments on the Food and Drug Administration (FDA) draft *Guidance for Industry Pharmacogenomic Data Submission* published in the November 4, 2003 *Federal Register*. The National Institutes of Health (NIH) strongly supports FDA's effort to address proactively the use of pharmacogenetic and pharmacogenomic (collectively referred to as PGx) strategies in drug development and utilization. The science and technology underlying PGx applications are evolving rapidly, and the draft guidance is a positive step aimed at promoting the application of PGx knowledge to improve public health while acknowledging that further research is still needed. In particular, we commend FDA for taking a creative approach to guidance in an important scientific arena, for the use of a "safe harbor" mechanism to encourage data submission, and for differentiating between research and regulatory uses of the data. While supportive of the draft Guidance in general, NIH has several specific suggestions for FDA's consideration.

Recommendations

It would be helpful if the Guidance could provide more specific guidance in several areas. First, the Guidance should be clearer about how the agency will decide when PGx data will be required for submissions of Investigational New Drug applications, Biologics License applications, and New Drug applications. Moreover, it would be helpful to clarify the criteria that will be used to determine what constitutes a "valid biomarker" and a "probable valid biomarker." As written, the draft Guidance simply states that the same rules that are applied to other biomarkers (e.g., serum) will be applied to PGx biomarkers. Since current data on the clinical significance of any given single nucleotide polymorphism (SNP) and the clinical validity of expression associations are lacking, this standard will be very difficult to meet. Further, given how many SNPs and expression patterns exist, carrying out the epidemiological research needed to determine their clinical significance will be daunting. To the extent possible, the Guidance should identify minimum acceptable standards of validation for PGx data and be updated frequently when new data warrants it. Without such minimums, sponsors and investigators are likely to avoid using PGx data or developing PGx tests for fear of undermining products in "susceptible" product development phases, where questions as to validity of the data can cause development delay or investor concern. Moreover, this could have public health implications if it delayed development and marketing of important PGx tests.

Second, the Guidance could be clearer and more specific regarding the definition of Voluntary Genomic Data Submissions (VGDS), the kind of data that should be submitted on a voluntary basis, how the voluntarily submitted data will be handled, stored, and distributed, and how and by whom such data will be evaluated. Will VGDS be reviewed in the context of other data submitted for regulatory purposes? Will they be associated with the submitter? In addition, the Guidance should be clearer about whether VGDS will ever be used for regulatory decisions. In particular, the Guidance states at line 311 and 498 that data submitted voluntarily will not be used for regulatory purposes. However, this assurance appears inconsistent with the statement at lines 312 -314 that "if additional information becomes available that renders the results required to be submitted . . . the sponsor must submit the data to the IND, BLA, or NDA." It would be helpful to clarify what "renders the results required to be submitted" means and how this determination will be made. Will these decisions be made on a case-by-case basis? Will they be made

unilaterally or through a consultation process with the sponsor? We would recommend that FDA develop a data template to guide data submission. Such a template would help submitters but might also be useful to the FDA as an efficient means for data aggregation and analysis. We would recommend that clear mechanisms for resolving disputes that might arise about the validity of a biomarker be in place and spelled out in the Guidance. In addition, we believe that if incentives beyond the safe harbor mechanism for voluntary data submission were provided, the Guidance might be more effective in achieving its intended purposes.

While NIH believes the Guidance will be beneficial overall, we are concerned that it might also have unintended negative effects on the willingness of sponsors to share PGx data. The FDA should clarify whether VGDS data will be considered confidential commercial information and, as such, not publicly accessible. If the data are considered confidential, we would urge FDA to work out a way to blind or aggregate the data so that it can be made available without compromising its proprietary status.

Pharmacogenomic data is developmentally sensitive. For example, since drug response expression patterns change throughout life, findings from one age group will not necessarily apply to another. As such, the Guidance should promote the submission of data from all age groups.

It would be helpful if the Guidance could also explain how pharmacogenomics may affect designation of orphan exclusivity for drugs. For example, if validated pharmacogenomic tests lead to the conclusion that certain small subsets of a population demonstrate selective response to a drug product, will each intended use be accorded orphan drug status?

We also believe that the goals of the Guidance could apply to proteomic data as well as PGx data. The field of proteomics is developing quickly, overlaps with pharmacogenomics in significant ways, and is equally in need of guidance about how it could and should be used in the regulatory process. We, therefore, urge FDA to incorporate provisions into the Guidance that could accommodate proteomic data submission or, if more appropriate, develop a separate Guidance for proteomics.

Conclusion

The NIH appreciates the opportunity to comment on the draft *Guidance for Industry Pharmacogenomic Data Submission* Guidance and commends FDA for developing an approach that is sufficiently in touch with cutting-edge science, yet flexible enough to accommodate further and perhaps unpredictable developments in the field.

We would welcome the opportunity to assist FDA in the further refinement of the Guidance, particularly determining what voluntary data would most be useful and the standards for determining the validity of PGx. The trans-NIH Biomedical Information Science and Technology Initiative (BISTI) Consortium and the Pharmacogenetics Research Network are two programs of particular relevance that could, respectively, provide a forum for further collaboration and serve as a model for data collection. Less structured opportunities for discussion and development of policies and practices concerning PGx data would be welcome, e.g., interaction with NIH scientific experts in the relevant research fields.

Iconix Pharmaceuticals Inc.

FDA Draft Guidance on Pharmacogenomic Data Submissions

Recommendations and Proposals

Please find enclosed the comments from Iconix Pharmaceuticals regarding the Draft Guidance for Industry on Pharmacogenomic Data Submissions. Our comments can be divided into three main categories: study design, data submission, and data interpretation, which are put into context in the accompanying flow diagram (see figure below). Note that these recommendations are for pharmacogenomic data in support of preclinical studies, and do not necessarily apply to clinical pharmacogenomic data. The comments are also focused on the generation, analysis and validation of 'omic' data.

1. Study Design Principles

Rigorous study design principles are a prerequisite to generating and effectively interpreting transcriptional data. Iconix recommends that a minimum of a biological triplicate per treatment (e.g. 3 animals per experimental and control group) be included in the study design. Experimental results have shown that this design parameter helps to control for biological variation and allows statistics to be used in the analysis. In addition, controls (e.g., vehicle-treated or sham-treated) should be processed within the same time frame as experimental samples in order to control for process drift. Since process drift is a Sponsor-dependent event, samples not processed within a relatively short time frame (e.g., less than a month) should be accompanied by sufficient quality control data to ensure data integrity. Iconix also recommends the use of the universal external RNA standard, when it becomes available, which should be processed in the same time frame as the submitted data in triplicate. This standard will aid in detecting quality control issues and process drift.

2. Pharmacogenomic Data Integrity and Quality.

(i) Data Integrity

Since FDA reviewers have a limited amount of time to review an IND submission, a rapid assessment of data integrity is essential. It is our recommendation that the Agency set standards for the minimum criteria to assure that the submitted data are of suitable quality for further analysis. The target format should include components of the MIAME standards, as well as supporting data typical of a peer-reviewed publication with some exceptions as noted below.

All RNA samples should be assessed for quality. For total RNA, metrics such as 28S/18S ratios are currently being used. This standard appears to be adequate for procedures employing total RNA. However this is not universally the case. In other processes, including the one used at Iconix, 'enriched' mRNA samples are used

effectively. Iconix has experience running over 15,000 microarray experiments and has empirically determined that the 28S/18S ratio is not an appropriate quality metric for this type of RNA sample. An alternative method to identify degraded samples is obtained using electropherograms generated on an Agilent Bioanalyzer. However the inclusion of electropherograms as part of a submission to the Agency is not recommended due to the difficulty of quickly interpreting these results. Rather an auditable statement from the sponsor that samples passed set criteria would be preferable.

(ii) Data Quality

It is also recommended that array quality be assessed for each submitted dataset. A basic set of metrics including median signal to background, average normalized background, log dynamic range, and mean raw signal across the array should calculated and used to determine quality. Comparing these values to historical data and/or to the external standards facilitates benchmarking the relative quality of the test arrays. It is important to note that although spike-in bacterial or yeast controls are used by many laboratories (including our own) in each array experiment, we do not consider them to be as robust as other metrics of hybridization performance. Therefore, we recommend against the use of the results of spike-in controls as a quality metric.

As a further valuable quality check, Iconix recommends the implementation of correlation analysis of the log signal intensity of each array hybridization experiment to several known control tissue references. A poor correlation to the respective tissue reference for a particular array experiment quickly identifies a poorly processed array and/or sample mishandling.

(iii) Data Submission

With regard to submission of data via a report format, Iconix recommends that line 458 of the Draft Guidance, "validation of gene expression by conventional assays..."; and line 462, "submission of electronic file containing raw images and scatter plots" be omitted as this data can not be readily generated or easily reviewed. For example, it is impractical to require northern analyses to accompany and support transcriptional changes measured on a microarray.

Each sponsor may have different preferences for array platform and analysis tools. As long as quality assessment metrics are minimally satisfied and data are submitted in a reproducible, interpretable format, the Agency should be accommodating to facilitate and not restrict the types of analysis or interpretation conducted by individual Sponsors who voluntarily submit pharmacogenomic data.

However for required preclinical pharmacogenomic data submission, it is recommended that commercial hardware and software used in data generation and analysis be revealed to the Agency.

3. The Importance of Context for Biomarker Validation

The contextual nature of data interpretation defines the level of validation of a biomarker. Iconix believes that greater detail is needed to describe the process of elevating a biomarker from "exploratory research pharmacogenomic data" to "probable valid biomarker" to "known valid biomarker" (Figure 1).

It is recommended that the categorization of the biomarker be established by (i) the quality of study design from which it came, (ii) thorough QC analysis (noted above), and (iii) an adequate level of interpretation and validation. It is recommended that for evaluation of the authenticity of a known valid biomarker, the Agency should have access to the entire gene expression dataset that led to the conclusion or at least to the minimum dataset required to reach the conclusion.

It is further recommended that an appointed body be responsible for judging the status of biomarkers. This is a role that could be performed by the IPRG if it were adequately supported.

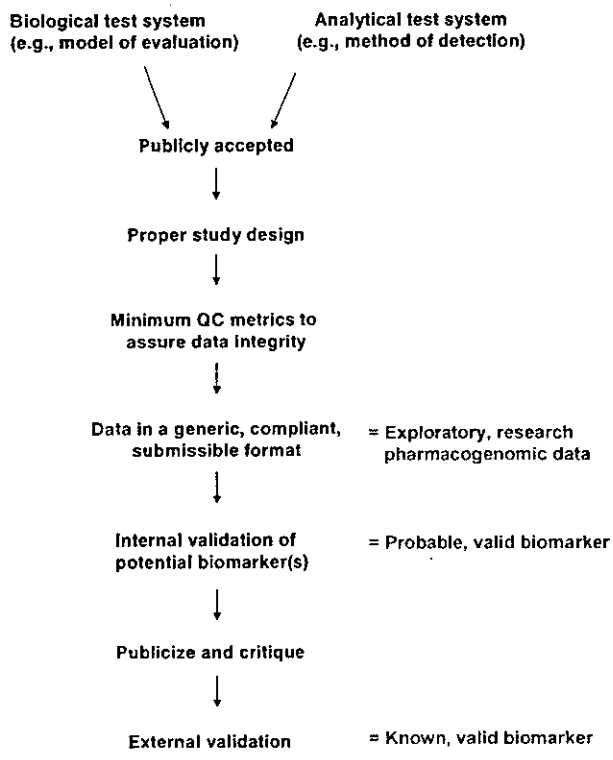


Figure 1: Flow process for development of a known, valid biomarker