Biomarker References for Pharmaceutical Statisticians

This document is intended as a useful reference list for pharmaceutical statisticians working with biomarkers. Articles are arranged into sections and briefly described so that the most relevant reference for a given topic can be found.

General Introductory Papers

Articles that appear in this section give a non technical and broader overview of the use, terminology and issues associated with the biomarkers.

Many of the references that appear throughout this entire document are taken from the following review paper, put together by the PSI Biomarker Special Interest Group. This may provide a useful general introduction to statistical issues in the use of biomarkers:

Jenkins M, Flynn A, Smart T, Harbron C, Sabin T, Ratnayake J, Delmar P, Herath A, Jarvis P, Matcham J, on behalf of the PSI Biomarker Special Interest Group, **A statistician's perspective on biomarkers in drug development**, Pharmaceutical Statistics, Volume 10, Issue 6, pages 494–507, November/December 2011, DOI: 10.1002/pst.532, http://onlinelibrary.wiley.com/doi/10.1002/pst.532/abstract

> A wide-ranging, high level review of statistical considerations in the use of biomarkers, including examples, terminology, practical challenges, handling large numbers of biomarkers, personalized medicine considerations, toxicity markers and qualification.

Examples of biomarkers:

These items provides the status-quo with respect to the regulatory (FDA) approvals of biomarkers for use in drug labels:

Table of pharmacogenomic biomarkers in drug labels. Available at: http://www.fda.gov/drugs/scienceresearch/researchareas/pharmacogenetics/ucm083378.html

Frueh F, Amur S, Mummaneni P, Epstein RS, Aubert RE, DeLuca TM, Verbrugge RR, Burckart GJ, Lesko LJ. **Pharmacogenomic biomarker information in drug labels approved by the United States food and drug administration: prevalence of related drug use.** *Pharmacotherapy* 2008; **28**(8):992–998.

http://onlinelibrary.wiley.com/doi/10.1592/phco.28.8.992/pdf

A review of 1200 drug labels reviewed for the years 1945–2005, 121 drug labels contained pharmacogenomic information based on a key word search and follow-up screening. Of those, 69 labels referred to human genomic biomarkers, and 52 referred to microbial genomic biomarkers.

Biomarker Definitions and introduction:

Atkinson AJ, *et al*. (Biomarkers Definitions Working Group). **Biomarkers and surrogate endpoints: preferred definitions and conceptual framework**. *Clinical Pharmacology and Therapeutics* 2001; **69**(3):89–95. http://www.nature.com/clpt/journal/v69/n3/abs/clpt200113a.html

Although dated and not statistically focused this article provides a coherent descriptions of definitions of biomarkers and if often referenced.

Dancey JE, *et al*. **Guidelines for the Development and Incorporation of Biomarker Studies in Early Clinical Trials of Novel Agents**. *Clinical Cancer Research* 2010; **16**:1745. http://clincancerres.aacrjournals.org/content/16/6/1745.full

This paper provides standard definitions and categories of biomarkers, and lists recommendations to sponsors and investigators on the identification and prioritization of biomarkers and assays, the coordination of activities for the development and use of assays, and for operational activities. It also contains some high level considerations surrounding the types of biomarker hypotheses that can be evaluated in phase 1 and 2 trials.

Hodgson D, Whittaker R, Herath A, Amakaye D, Clack G. **Biomarkers in oncology drug development.** *Molecular Oncology* 2009; **3**(1):24–32. http://www.elsevierscitech.com/pdfs/molonc/10.biobankinginoncology.pdf

This review discusses the properties of biological sample based efficacy measurements and their implementation in oncology drug development, including points to consider and examples.

Statistical Reflections

The articles listed in this section have a more statistical focus. These are good overall reads and provide an opportunity to think more about the utility of biomarkers and also issues that might be inherent to biomarker discovery and subsequent use.

Carroll KJ. **Biomarkers in drug development: friend or foe? A personal reflection gained working within oncology**. *Pharmaceutical Statistics* 2007; **6**(4):253–260. http://onlinelibrary.wiley.com/doi/10.1002/pst.269/pdf

> A very good overview of the use and misuse of biomarkers in drug development, it points out many of the pitfalls and statistical issues. The paper explains why biomarkers may not be the answer to drug development in terms of the usual perceived advantages of de-risking the program and making it quicker and cheaper.

George S. **Statistical Issues in Translational Cancer Research**. *Clinical Cancer Research* 2008;14(19) http://clincancerres.aacrjournals.org/cgi/pmidlookup?view=long&pmid=18829473

> A general reflection on predictive models and biomarkers and why there may be some differences when using these in a more exploratory fashion compared to the thinking for a larger confirmatory clinical trial.

Veneis P, McMichael AJ. **Bias and confounding in molecular epidemiological studies: special considerations**. *Carcinogenesis* 1998; 19(12):2063–2067. http://carcin.oxfordjournals.org/content/19/12/2063.full.pdf+html

> This paper considers the possible causes of bias and confounding in epidemiological studies. Important considerations even if a randomized clinical trial is used. Biomarkers are considered with both the benefits and issues in different situations.

Ioannidis JPA, Panagiotouo OA. **Comparison of effect sizes associated with biomarkers reported in highly cited individual articles and in subsequent meta-analyses**. *Journal of the American Medical Association* 2011; **305**(21):2200–2210. http://jama.jamanetwork.com/article.aspx?articleid=900417

> Insightful article of propagation of effect sizes of biomarkers from the early discovery (small sized studies) to eventual populations.

Validation and Qualification

The first two papers in this section consider fit-for-purpose analytical validation of biomarkers, while the remaining papers consider how robust signatures and predictive models are to the individual teams developing them.

Lee JW, *et al*. Fit**-for-purpose method development biomarker and validation for successful biomarker measurement**. *Pharmaceutical Research* 2006; **23**(2):312–328. http://rd.springer.com/article/10.1007%2Fs11095-005-9045-3

> This detailed paper covers the issues around validation of lab biomarkers making suggestions on how the validation should be fit for purpose and this will depend on the planed use of the biomarker

Joel T Dudley, Robert Tibshirani, Tarangini Deshpande & Atul J Butte, **Disease signatures are robust across tissues and experiments**, Molecular Systems Biology 5 Article number: 307., doi:10.1038/msb.2009.66, http://www.nature.com/msb/journal/v5/n1/full/msb200966.html

> Meta-analyses were applied to publically available microarray data that included both normal control and diseased state. This consisted of 429 experiments, representing 238 diseases and 122 tissues from a total of 8435 arrays. Concordance found between experiments measuring the same disease condition suggesting disease signatures may be robust across experiments. It was found that the molecular signature of disease across time is more prominent than the signature of tissue expression across disease.

Shi L. *et al*. **The MicroArray Quality Control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models**. *Nature Biotechnology* 2010; **28**(8):827–838. http://www.nature.com/nbt/journal/v28/n8/full/nbt.1665.html

> Reliability of predictions from gene expression data was assessed with 36 independent teams analyzing 6 microarray datasets to generate predictive models. The performance of the models depended on the endpoint and the team proficiency. Different approaches generated models of similar performance. Good modeling practice guidelines were established

Safety Biomarkers

Sistare FD et al. **Towards consensus practices to qualify safety biomarkers for**

use in early drug development. *Nature Biotechnology* 2010; **28**:446–454. http://www.nature.com/nbt/journal/v28/n5/full/nbt.1634.html

> A consortium was formed to consider the validity of safety biomarkers that could be used in translation from pre-clinical to clinical and in early phases of development. Guidelines and considerations for some aspects of fit-for-purpose qualification of safety biomarkers were given, emphasizing the need to consider what is required as early as possible.

Warnock D, Peck C. **A roadmap for biomarker qualification**. *Nature Biotechnology* 2010; **28**(5):444–445. http://www.nature.com/nbt/journal/v28/n5/full/nbt0510-444.html

Dieterle F, Sistare F, Goodsaid F *et al*. **Renal biomarker qualification submission: a dialog between the FDA-EMEA and Predictive Safety Testing Consortium**. *Nature Biotechnology* 2010; **28**(5):455–462. http://www.nature.com/nbt/journal/v28/n5/full/nbt.1625.html

> Collaborative efforts between pharmaceutical companies, regulatory agencies and academia to qualify biomarkers for kidney toxicity - provides a model for investigating and identifying reliable safety markers for preclinical applications

Statistical Methodology

The articles in this section introduce some of the statistical methods often used in biomarker based studies.

Stephen W. Looney, Joseph L. Hagan, 4 **Statistical Methods for Assessing Biomarkers and Analyzing Biomarker Data**, Handbook of Statistics,Volume 27, 2007, Pages 109–147 http://www.sciencedirect.com/science/article/pii/S016971610727004X

> In this chapter, the authors provide a comprehensive appraisal of a vast array of statistical methods that could be used when analyzing biomarker data. A useful first read in the area of statistics methods applied to biomarker data.

David P. Lovell., **Commentary: statistics for biomarkers**, Biomarkers. 2012 May;17(3):193-200. doi: 10.3109/1354750X.2012.656287. http://www.ncbi.nlm.nih.gov/pubmed/22332747

> A short commentary that discusses requirements for the reporting of statistical analyses of biomarkers in papers submitted to 'Biomarkers' journal. Paper

provides top level overview of good statistical practice rather than given details of methodology that could be applied.

Altman DG, Bland JM. **Diagnostic tests 2: predictive values**. *British Medical Journal* 1994; **309**:102. http://www.bmj.com/content/309/6947/102.1

Gentle introduction to diagnostic tests (this is the part two introducting positive/negative predictive values and their cousins). This is part of a series of Statistics Notes by Bland and Altman in the BMJ (see http://www-users.york.ac.uk/~mb55/pubs/pbstnote.htm for more notes).

Martin Bland's page on measurement studies is also very relevant when considering comparisons between scoring methods (see http://www-users.york.ac.uk/~mb55/meas/meas.htm)

Long Q, *et al*. **Robust statistical methods for analysis of biomarkers measured with batch/experiment-specific errors**. *Statistics in Medicine* 2010; **29**:361–370. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3177604/

In this paper, the authors provide approaches to (1) modeling the association between an outcome and an explanatory variable that was measured with batch/experiment-specific errors and (2) evaluating the diagnostic/predictive accuracy of such an explanatory variable when the outcome of interest is a disease status. Two robust methods (RMs) are proposed that do not rely on assumptions to be made on the structure and distribution of measurement errors.

Benjamini Y, Hochberg Y. **Controlling the false discovery rate: a practical and powerful approach to multiple testing**. *Journal of the Royal Statistical Society, Series B* 1995; **57**(1):289–300. http://www.jstor.org/discover/10.2307/2346101?uid=3738032&uid=2129&uid=2&uid=70&uid=4&sid=21102611851747

Storey JD. **A direct approach to false discovery rates**. *Journal of the Royal Statistical Society, Series B* 2002; **64**:479–498. http://onlinelibrary.wiley.com/doi/10.1111/1467-9868.00346/pdf

Two papers describing the false discovery rate (FDR), an alternative solution to the issue of multiplicity when dealing with high dimensional technologies. Rather than controlling the Type I error as in standard approaches for multiplicity which would impose an unrealistically severe penalty with potentially thousands of variables, the FDR accepts that there will be false positives and estimates the proportion of false positives in a set of results.

High dimensional data

Although not exclusive to biomarkers, many omic technologies produce high dimensional datasets which require specific multivariate methods to model them. Below are a variety of methods for fitting high-dimensional prediction models. These use a variety of strategies to avoid the issues of over-fitting when the number of predictor variables approaches or exceeds the number of observations. These methods can be used both to generate predictive models, and also for variable selection reducing a large set of candidate variables to a smaller set for more detailed analysis.

Parry RM, Jones W, Stokes TH, Phan JH, Moffitt RA, Fang H, Shi L, Oberthuer A, Fischer M, Tong W, Wang MD. **k-Nearest neighbor models for microarray gene expression analysis and clinical outcome prediction**. *The Pharmacogenomics Journal* 2010; **10**:292–309.
http://www.nature.com/tpj/journal/v10/n4/full/tpj201056a.html

> High-Dimensional Prediction using Nearest Neighbours, a proximity based algorithm.

Breiman L. **Random Forests**. *Machine Learning* 2001; **45**(1):5–32.
http://rd.springer.com/article/10.1023%2FA%3A1010933404324

> High-Dimensional Prediction using Random Forests, a tree based methodology.

Cortes C, Vapnik V. **Support-Vector Networks**. *Machine Learning* 1995; **20**:273–297.
http://rd.springer.com/article/10.1007%2FBF00994018

> High-Dimensional Prediction using an algorithm giving maximal separation between groups

J.H. Friedman. **Greedy Function Approximation: A Gradient Boosting Machine** *Annals of Statistics* 2001; **29**(5):1189-1232.
http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.29.9093

> High-Dimensional Prediction using Gradient Boosting Machine, a tree based methodology.

Friedman, J., Hastie, T. and Tibshirani, R. **Regularization Paths for Generalized Linear Models via Coordinate Descent**, *Journal of Statistical Software,* 2008; **33**(1), 1-22 http://www.jstatsoft.org/v33/i01/paper

> High-Dimensional Prediction using Elastic Nets, a regularized form of

regression, including Lasso and Ridge Regression as special cases.

ROC curves

Jafarzadeha SR, Johnson WO, Utts JM, Gardner IA. **Bayesian estimation of the receiver operating characteristic curve for a diagnostic test with a limit of detection in the absence of a gold standard.** *Statistics in Medicine* 2010; **29**:2090–2106. http://onlinelibrary.wiley.com/doi/10.1002/sim.3975/pdf

> This paper focuses methodology taking account of data below/above limit of detectyion when determing receiver operating characteristic (ROC) curve to evaluate the discriminatory ability of a biomarker. Ignoring the scores that are beyond the limit of detection of a test leads to a biased assessment of its discriminatory ability. The authors present a Bayesian approach for the estimation of the ROC curve and its AUC.

Hsieh H-N, Su H-Y, Zhou X-H. **Interval estimation for the difference in paired areas under the ROC curves in the absence of a gold standard test**. *Statistics in Medicine* 2009; **28**:3108–3123.
http://onlinelibrary.wiley.com/doi/10.1002/sim.3661/pdf

Joseph L, Gyorkos TW, Coupal L. **Bayesian Estimation of Disease Prevalence and the Parameters of Diagnostic Tests in the Absence of a Gold Standard**. *rican Journal of Epidemiology* 1995; **141**(3):263–272.

http://aje.oxfordjournals.org/cgi/pmidlookup?view=long&pmid=7840100

Choi Y-K, Johnson WO, Collins MT, Gardner IA. **Bayesian Inference for Receiver Operating Characteristic Curves in the Absence of a Gold Standard**. *Journal of Agricultural, Biological, and Environmental Statistics* 2006; **11**(2):210–229.

http://rd.springer.com/article/10.1198%2F108571106X110883

> These three papers presents methodology that can be used when determining ROC curves when a Gold Standard test may sometimes be too expensive or infeasible. For example, in many medical research studies, the true disease status of the subjects may remain unknown. In the first paper Maximum Likelihood-based procedure using the expectation-maximization (EM) algorithm in conjunction with a bootstrap method for the construction of confidence intervals for the difference in paired AUCs in the absence of a gold standard test. The latter two papers use Bayesian approaches to make inferences about ROC curves.

Clinical Study Designs

The articles in this section consider specific clinical trial designs that can be used to demonstrate the clinical utility of a biomarker-based classifier.

Gary J. Kelloff1 & Caroline C. Sigman, **Cancer biomarkers: selecting the right drug for the right patient**, Nature Reviews Drug Discovery 11, 201-214 (March 2012) | doi:10.1038/nrd3651,
http://www.nature.com/nrd/journal/v11/n3/abs/nrd3651.html

> A very useful paper covering a range of innovative biomarker based study designs, and methodologies for defining, evaluating and using biomarker classifiers. Designs discussed include randomization or stratification by biomarker status, enrichment designs, and various other adaptive approaches including response-adaptive, bayesian adaptive randomization, and the adaptive signature design.

Freidlin B, McShane LM, Korn EL. **Randomized clinical trials with biomarkers: design issues**. *Journal of the National Cancer Institute* 2010; **102**:152–160.
http://jnci.oxfordjournals.org/content/102/3/152

> This paper provides an in-depth comparison of the strengths and weaknesses of some of the commonly used biomarker designs (including Biomarker stratified, Enrichment and biomarker strategy designs).

Simon R. **Advances in clinical trial designs for predictive biomarker discovery and validation**. *Current Breast Cancer Reports* 2009; **1**:216–221.
http://rd.springer.com/article/10.1007%2Fs12609-009-0030-4

> Another article providing a high level overview of recent developments/publications in biomarker study designs.

Buyse M, Michiels S, Grothey A, Matheson A, De Gramont A. **Integrating biomarkers in clinical trials**. *Expert Review of Molecular Diagnostics* 2011; **11**(2):171–182.
http://www.expert-reviews.com/doi/abs/10.1586/erm.10.120

> Among many aspects of biomarkers in clinical trials, The paper discusses 9 different trial designs useful under different experimental setups: (A) Discordant risk randomized design (B) intermediate-risk randomized design (C) randomize-all design (D) interaction or biomarker-stratified design (E) biomarker-strategy design with standard control; (F) biomarker-strategy design with randomized control (G) Bayesian adaptive Phase II design (P1, P2 and so

on: probabilities of allocating) with examples of their usage.

Hoering A, LeBlanc M, Crowley JJ. **Randomized phase III clinical trial designs for targeted agents**. *Clinical Cancer Research* 2008; **14**(14):4358–4367. http://clincancerres.aacrjournals.org/content/14/14/4358

This article evaluates the effectiveness of the randomize all, the targeted, and the strategy phase III trial designs under: the presence of a prognostic marker, presence of a predictive marker and the absence of a valid marker. Only biomarkers of continuous in nature are considered. It also investigates the performance of several test statistics for the different trial designs as a function of the marker distribution and the marker cutoff. The performance is evaluated as a function of the cut point, the number of patients screened, and the number of patients randomized to obtain a certain power and significance for the various test statistics.

Some Selected Adaptive Designs

Wang S-J, O'Neill RT, Hung HMJ. **Approaches to Evaluation of Treatment Effect in Randomised Clinical Trials with Genomic Subset**. *Pharmaceutical Statistics* 2007; **6**(3):227–244. http://onlinelibrary.wiley.com/doi/10.1002/pst.300/pdf

Paper discusses adaptive & non-adaptive clinical trial designs to evaluate treatment effects related to genomic profiles or genomic composite biomarkers. Sample size, power and multiplicity adjustments are discussed under the two trial designs.

Freidlin B, Simon R. **Adaptive signature design: an adaptive clinical trial design for generating and prospectively testing a gene expression signature for Sensitive Patients**. *Clinical Cancer Research* 2005; **11**(21):7872–7878. http://clincancerres.aacrjournals.org/content/11/21/7872

This paper proposes a new adaptive design for randomized clinical trials of targeted agents in settings where an assay or signature that identifies sensitive patients is not available at the outset of the study. The design combines prospective development of a gene expression–based classifier to select sensitive patients with a properly powered test for overall effect. Simulated results from a traditional design and adaptive design are discussed.

Jiang W, Freidlin B, Simon R. **Biomarker adaptive threshold design: a Procedure for Evaluating Treatment with Possible Biomarker-Defined Subset Effect**.

Journal of the National Cancer Institute 2007; **99**(13):1036–1043. http://jnci.oxfordjournals.org/content/99/13/1036

> The paper presents simulated results to evaluate the performance of the adaptive design, relative to the more traditional design, under two conditions: (a) when the proportion of patients sensitive to the new drug is low, the adaptive design substantially reduces the chance of false rejection of effective new treatments. (b) when the new treatment is broadly effective, the adaptive design has power to detect the overall effect similar to the traditional design. Formulas are provided to determine the situations in which the new design is advantageous.

Surrogate Endpoints

Although the interest in surrogate endpoints has waned over recent years (as the difficulty in making the case for surrogacy became clear), a knowledge of the methods involved can be useful when discussing what can (and cannot) really be concluded as a result of biomarker based clinical studies. The methods below can also be employed when considering likely effect sizes when translating between different endpoints.

Evidence for surrogacy:

Fleming TR, DeMets DL. **Surrogate endpoints in clinical trials: are we being misled?** *Annals of Internal Medicine* 1996; **125**:605–613. http://annals.org/article.aspx?articleid=710042

Fleming T. **Surrogate endpoints and FDA's accelerated approval process**. *Health Affairs* 2005; **24**(1):67–78. http://content.healthaffairs.org/content/24/1/67

> These papers by Fleming explain the difficulties in establishing surrogacy.

Baker S, Kramer B. **A perfect correlate does not a surrogate make**. *BMC Medical Research Methodology* 2003; **3**:16. http://www.biomedcentral.com/1471-2288/3/16

> A short, simple paper, but one which can be useful to be aware of when explaining surrogacy to colleagues. A simple figure is used to explain why correlation between endpoints at an individual patient level does not imply surrogacy.

Buyse et al. **Biomarkers and surrogate end points—the challenge of statistical validation**. *Nature Reviews Clinical Oncology* 2010; **7**:309–317. http://www.nature.com/nrclinonc/journal/v7/n6/full/nrclinonc.2010.43.html

A readable paper discussing practical issues with biomarker validation in plain language. The first half of the paper covers prognostic and predictive biomarkers and the second half covers surrogate endpoints.

Overview and case study:

Weir C, Walley R. **Statistical evaluation of biomarkers as surrogate endpoints: a literature review**. *Statistics in Medicine* 2006; **25**:183–203.
http://onlinelibrary.wiley.com/doi/10.1002/sim.2319/pdf

A useful and wide-ranging review paper starting from the earlier work of Prentice and Freedman, moving right up to more recent meta-analytic approaches. Includes technical details and many references.

Qian Shi, Sargent D. **Meta-analysis for the evaluation of surrogate endpoints in cancer clinical trials.** *Int J Clin Oncol* 2009; 14: 102-111
http://rd.springer.com/article/10.1007%2Fs10147-009-0885-4#page-1

A review of surrogate endpoint meta-analyses in cancer, the drugs and approvals they supported and methods used. A useful reference if considering the evidence available for surrogates in this therapy area.

Statistical methods:

Prentice RL. **Surrogate endpoints in clinical trials: definition and operational criteria.** *Statistics in Medicine* 1989; **8**:431–440.
http://onlinelibrary.wiley.com/doi/10.1002/sim.4780080407/abstract

Freedman LS, Graubard BI, Schatzkin A. **Statistical validation of intermediate endpoints for chronic diseases.** *Statistics in Medicine* 1992; **11**:167–178.
http://onlinelibrary.wiley.com/doi/10.1002/sim.4780110204/pdf

The first major attempts to define surrogate endpoints and the criteria that they should meet. Freedman's work moved the thinking towards the proportion of variance explained rather than statistical tests. Both are perhaps more clearly explained by the papers by Buyse and Burzykowski.

Buyse M, Molenberghs G. **Criteria for the validation of surrogate endpoints in randomized experiments**. *Biometrics* 1998; **54**:1014–1029.
http://www.jstor.org/discover/10.2307/2533853?uid=3738032&uid=2129&uid=2

&uid=70&uid=4&sid=21102611851747

Buyse M, Molenberghs G, Burzykowski T, Renard D, Geys H. **The validation of surrogate endpoints in meta-analyses of randomized experiments**. *Biostatistics* 2000; **1**(1):49–67. http://biostatistics.oxfordjournals.org/content/1/1/49

The first of a series of papers by Buyse, Molenburghs and Burzykowski introducing meta-analytic methods for the evaluation of surrogate endpoints. This was developed for various types of clinical endpoint (see review papers) and meta-analyses are now the methods most commonly employed. The 2000 paper contains two motivating examples.

Burzykowski T, Buyse M. **Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation**. *Pharmaceutical Statistics* 2006; **5**:173–186. http://onlinelibrary.wiley.com/doi/10.1002/pst.207/pdf

An important technical paper in which the authors build on their earlier meta-analytic methods to introduce the surrogate threshold effect - defined as the minimum treatment effect on the surrogate necessary to predict a non-zero effect on the true endpoint. Variability and prediction limits are considered as well as just measures in terms of a single coefficient.

Li Y, Taylor J. **Predicting treatment effects using biomarker data in a meta-analysis of clinical trials**. *Statistics in Medicine* 2010; **29**:1875–1889. http://onlinelibrary.wiley.com/doi/10.1002/sim.3931/pdf

A review and simulation study of several meta-analytic methods as well as Buyse et al.

Reporting of Biomarker Studies

The REMARK guidelines provide a checklist when reporting tumour biomarker based studies (although the principles would apply to other biomarker studies as well), particularly when writing a publication. Specific worked examples are given to these guidelines in the second paper:

McShane L, Altman A, Sauerbrei W, Taube S, Gion M, Clark G for the Statistics Subcommittee of the NCI–EORTC Working Group on Cancer Diagnostics, **REporting recommendations for tumor MARKer prognostic studies (REMARK),** *Nature*

Clinical Practice, 2(8), 2005
http://www.nature.com/nrclinonc/journal/v2/n8/full/ncponc0252.html

Altman D, McShane L, Sauerbrei S, Taube S, **Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): Explanation and Elaboration**, *PLoS Medicine* 2012, 9 (5) , http://www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.1001216

Regulatory Guidance

A knowledge of the following regulatory concept papers and guidance documents may be informative.

FDA. **Drug diagnostics co-development concept paper** (Draft, August 2005). Available at: http://www.fda.gov/downloads/Drugs/ScienceResearch/ResearchAreas/Pharmacogenetics/UCM116689.pdf

> This FDA document gives a good general overview of the steps involved in developing a drug and diagnostic simultaneously for the situation where a diagnostic test will be part of the clinical use of the drug. Important considerations include analytical test validation, clinical test validation and demonstration of clinical utility. The addenda also give some guidance for considering the performance of a diagnostic test.

FDA. Draft guidance for industry and food and drug administration staff – **In vitro companion diagnostic devices**, July 2011. Available at: http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm262292.html

> This FDA guidance is for sponsors who are planning to develop a therapeutic product that *depends on* the use of an in vitro (IVD) companion diagnostic device. Worth some awareness, but not statistically detailed. IVDs provide information that is essential for the safe and effective use of a corresponding therapeutic product and so are required to meet premarketing authorization.

FDA. Guidance for Industry: **Enrichment Strategies for Clinical Trials to Support Approval of Human Drugs and Biological Products,** Draft guidance, December 2012
http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM332181.pdf

> A useful read if considering a targeted development program where a biomarker defined subset of patients may respond better to the investigational treatment. Contains discussion on which populations should be studied in different situations and potential trial designs.

FDA. Guidance for Industry: **Qualification Process for Drug Development Tools**, October 2010. Available at:
http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM230597.pdf

Qualification of Novel methodologies for Drug Development: Guidance to applicants, Jan 2009. Available at:
http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2009/10/WC500004201.pdf

> These papers outline the qualification approaches introduced by the FDA and EMA that can be used to gain a regulatory opinion on the suitability of a biomarker for a given use. These are voluntary processes. For example for the FDA Once a drug development tool (DDT) has been qualified by the centre of drug evaluation and research (CDER) for a specific context of use, the DDT can be used in one or more drug development without the need to reconfirm the DDT's utility. The following links provides some examples of biomarkers qualification submissions to the FDA and EMA
> http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/ucm284076.html
>
> http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/document_listing/document_listing_000319.jsp&mid=WC0b01ac0580022bb0

European Medicines Agency. **Evaluation of Medicines for Human Use Innovative Drug Development Approaches** Final Report from the EMEA/CHMP-Think-Tank Group on Innovative Drug Development, 22 March 2007. Available at:
http://www.ema.europa.eu/docs/en_GB/document_library/Other/2009/10/WC500004913.pdf

> A high level view from regulators on a variety of issues. Biomarkers and surrogate endpoints are touched upon.

Genomics

FDA Guidance for Industry, Clinical **Pharmacogenomics: Premarket Evaluation in Early-Phase Clinical Studies and Recommendations for Labeling**, January 2013
http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM337169.pdf

FDA Guidance for Industry, **Pharmacogenomic data submissions** 2005

http://www.fda.gov/downloads/RegulatoryInformation/Guidances/ucm126957.pdf

ICH guideline **E16** on **genomic biomarkers** related to drug response: context, structure and format of qualification submissions:
http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2010/09/WC500097060.pdf

> Specific points to consider for the use of pharmacogenomics in submissions. The first paper contains some discussion on trial design and statistical considerations. FDA guidance covers how DNA variation can affect the PK and PD efficacy or safety in phase I and II stuides.