

AMGEN[®]

Pioneering science delivers vital medicines™



Assessing Similarity to Existing Drugs to Decide Whether to Continue Drug Development

Robert A. Parker, Amgen, Inc.

Joseph G. Ibrahim, University of North Carolina

Disclaimers (1)

- **The views expressed here are solely those of the authors and do not represent the views of Amgen, Inc.**

Disclaimers (2)

This talk is intended to present the concepts – it is **not** intended to be a detailed “how to do it” talk

Disclaimers (3)

- **Amgen has recently filed a BLA for an osteoporosis treatment***
- **Although osteoporosis is mentioned in several places as an illustration of how the approach could be applied, the actual Phase 3 decisions were made years before I arrived at Amgen and nothing in this talk actually presents any Amgen data**

*Amgen press release, available at http://wwwext.amgen.com/media/media_pr_detail.jsp?releaseID=1238183

Outline of Talk

- **Background**
- **Conventional Statistical Approach: Probability of Success**
- **Alternative Approach: Classification**

Background (1)

- **Drug development is expensive: it can cost > 1 billion dollars to get a drug to market**
- **Drug development costs increase dramatically in later stages: often increasing by an order of magnitude when going from Phase 1 to Phase 2 to Phase 3 studies**
 - a recent Amgen submission contained six Phase 3 studies with over 11,000 patients*
- **Therefore, one does not lightly go into a Phase 3 development program**

*Amgen press release, available at http://wwwext.amgen.com/media/media_pr_detail.jsp?releaseID=1238183

Background (2)

- **Amgen systematically reviews compounds multiple times during the development process to decide whether development should continue. Decisions can be**
 - GO to the next stage of development
 - NO-GO; discontinue development of this compound
 - NEED MORE INFORMATION
- **A major goal in drug development is to make NO-GO decisions as soon as possible**
 - avoids exposing patients to unnecessary risks
 - allows resources to be focused on drugs most likely to benefit patients
- **The problem is how to make these decisions wisely**

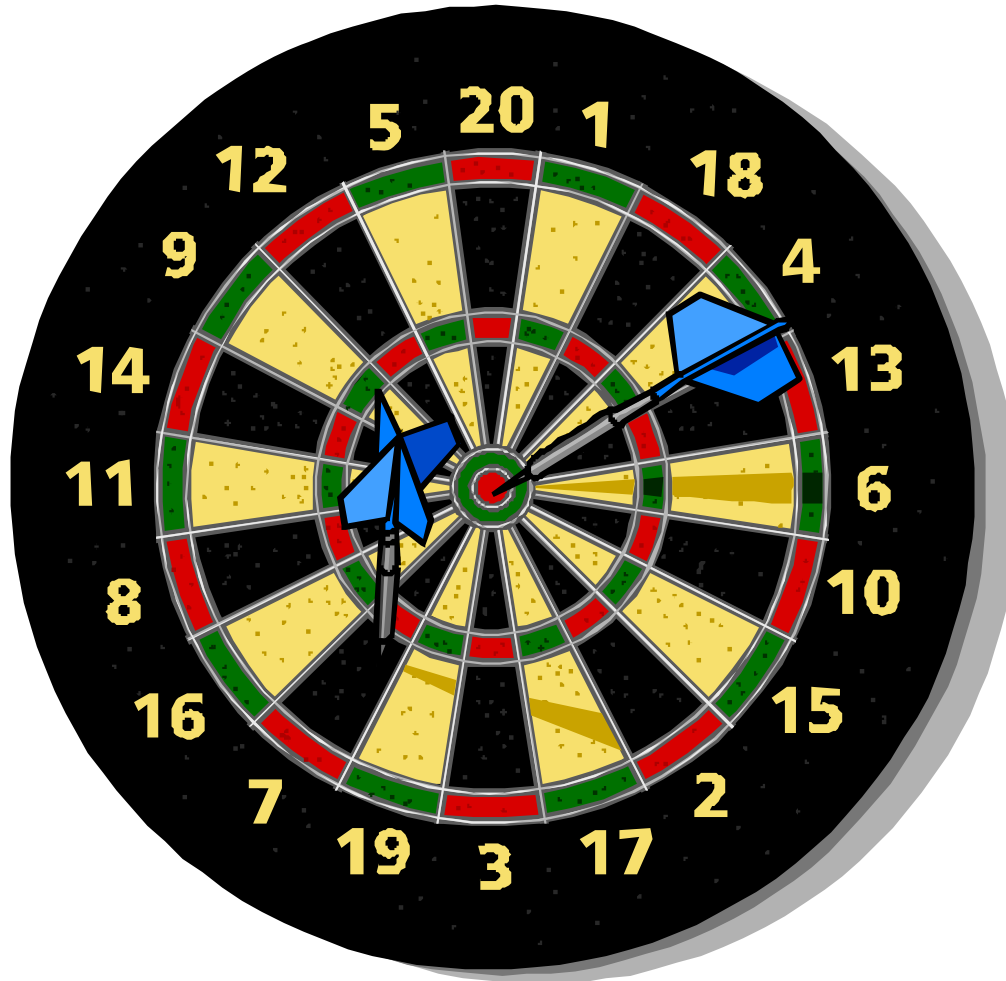
Traditional Methods (1)



Traditional Methods (2)



Traditional Methods (3)



Conventional Statistical Approach: Probability of Success

Probability of a Successful Study (1)

- **There are statistical methods to attempt to help in this decision. A traditional method would be to try to predict the chance that the compound is likely to be successful in the next stage(s)**
- **This could be done using a range of approaches from a relatively simply to a very complicated model**

Probability of a Successful Study (2)

- **A relatively simple approach would be to model the probability that the compound would meet prespecified criteria for Phase 3, based on the observed results of Phase 2**
 - **example would be the probability of success concept,* which incorporates both the probability that the result would be statistically significant and that the treatment effect would meet a minimum criterion for clinical importance**
 - **in the simplest case, one could use the effectiveness (or a fraction of the effectiveness to be conservative) from the Phase 2 study as a point estimate**

*Chuang-Stein, C. Sample size and the probability of a successful trial, Pharm Stat 2006;5;305-309.

Probability of a Successful Study (3)

- **This model could be expanded in many ways to incorporate uncertainty in the prediction:**
 - **reliability in the Phase 2 results treating results of Phase 2 as an observation from a distribution**
 - **population changes from Phase 2 to Phase 3**
 - **incorporation of a model for the observation that Phase 2 results tend to overestimate Phase 3 effectiveness**
 - **potential for safety issues to occur with the compound**

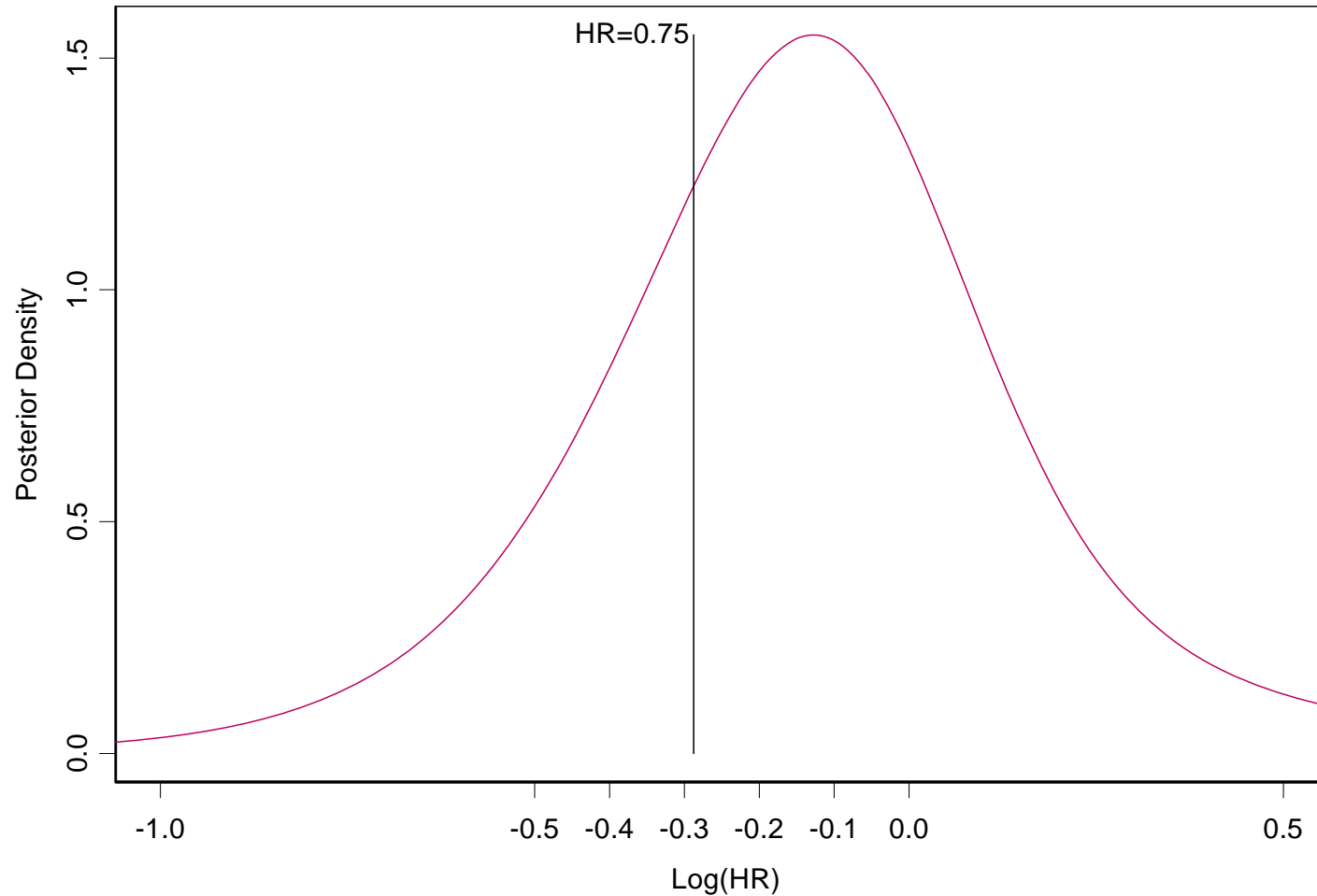
Probability of a Successful Study (4)

Example:

- time to event endpoint
- HR < 1 is benefit
- clinically very important benefit is HR < 0.75

Probability of a Successful Study (5)

Predicted P3 Log(HR)



Probability of a Successful Study (6)

- Can then use the distribution of $\log(\text{HR})$ with the proposed study design to determine the probability of success in Phase 3 (statistically significant and estimate is < 0.75)

$$\int_{\Theta} \Pr(\text{HR} < 0.75 \mid \theta) Pwr(\theta) PD(\theta) d\Theta$$

- $\Pr(\text{HR} < 0.75 \mid \theta)$: probability $\text{HR} < 0.75$, given θ
- $Pwr(\theta)$: power of the proposed study given θ
- $PD(\theta)$: prior distribution for θ

Probability of a Successful Study (7)

- However, an estimated HR=0.76, although not quite as big an effect as HR=0.75, is still likely to be of considerable value to patients
- So approach can be expanded to estimate clinical benefit of compound

$$\int CB(Ob sHR) Pr(Ob sHR | \theta) Pwr(\theta) PD(\theta) d\Theta$$

- ⊖
 - **CB(Ob sHR)**: clinical benefit of the observed HR
 - **Pr(Ob sHR | θ)**: probability of observing a specific HR, given θ
 - **Pwr(θ)**: power of the proposed study given θ
 - **PD(θ)**: prior distribution for θ

Alternative Approach: Classification

Classification (1)

- **Instead of trying to predict the probability of success using internally defined standards, focus on whether the compound looks like other successful drugs for the indication**
- **We know what a clinically important drug looks like (in terms of efficacy, side effect profiles, etc.) for many indications:**
 - osteoporosis
 - rheumatoid arthritis
 - hypertension
 - diabetes

Classification (2)

- We also know what would not be clinically useful, and what would be a marginally useful drug for these indications
- The **idea**, then, is to classify the results we observe with our new compound into one of these three groups, and make the decision based on the classification of the drug

Classification (3): Example

- **For osteoporosis, we can identify clinically important, marginal, and not clinically useful drugs**
- **In osteoporosis, the clinically important efficacy endpoint is fracture reduction**
- **There is a lot of evidence that a surrogate endpoint – change in bone mineral density (BMD) – is predictive of fracture reduction in larger studies, so we could use this short-term endpoint before going to larger Phase 3 fracture studies**

Classification (4): Example

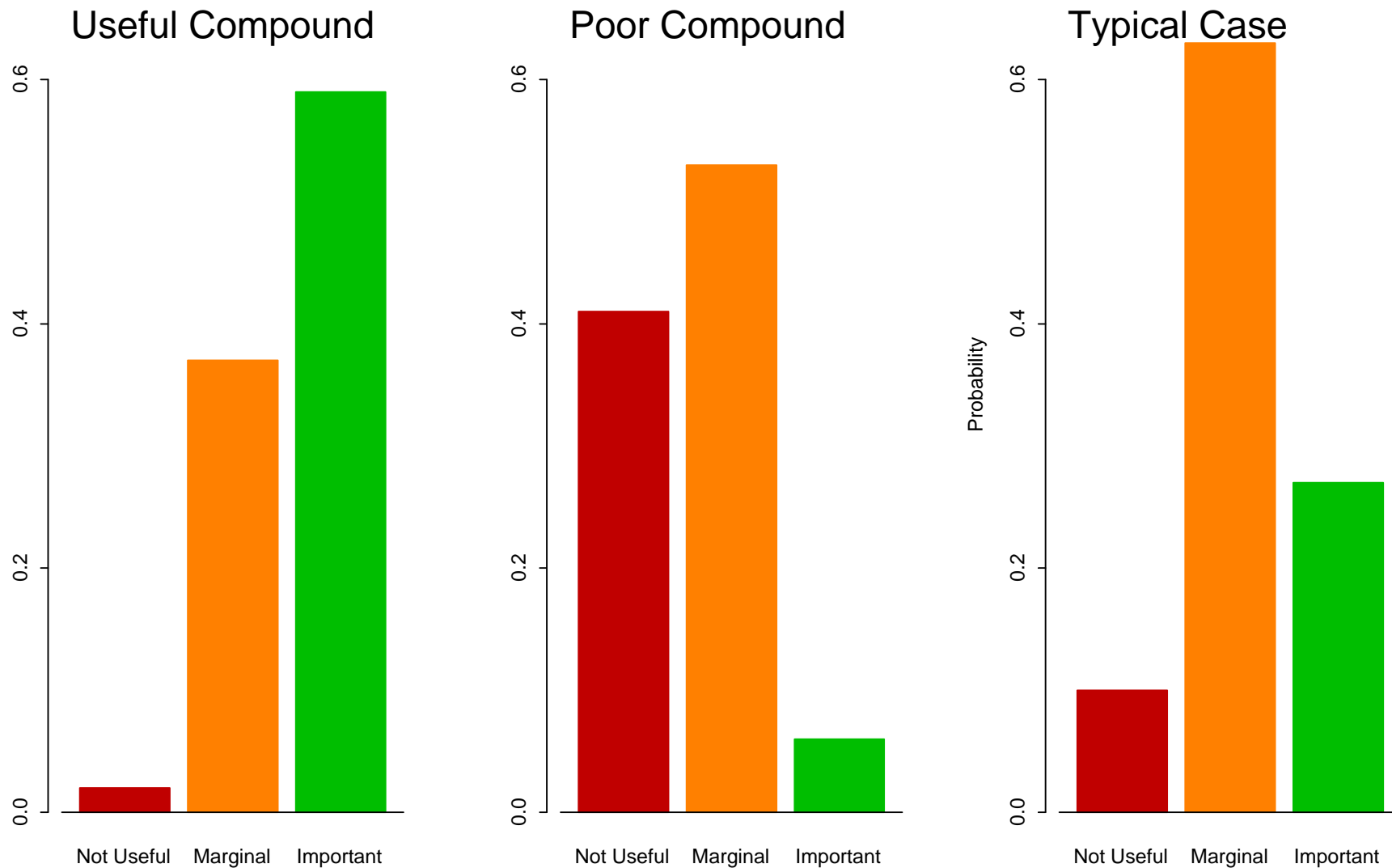
- **For efficacy, a clinically important compound* would raise bone mineral density (BMD):**
 - spine >8%
 - total hip >5%
- **For efficacy, a marginal drug would raise BMD less:**
 - spine 4-6%
 - total hip 2-3%
- **For efficacy, a not clinically useful drug would raise BMD little:**
 - spine 2-3%
 - total hip 0-1%

*Based on Fosamax® product information, 3 year results, retrieved 07 January 2009

Classification (5)

- The critical idea is that **the problem is changed** to one of predicting how likely it is that this new compound belongs to one of three classes:
 - clinically important ==> GO decision
 - marginal ==> ????
 - not clinically useful ==> NO-GO decision

Example of the Types of Results Provided



Conceptual Model: Components

- **drug class prior:** what "clinically important," "marginal," and "not clinically useful" means
 - almost always for multiple endpoints
 - potentially multiple separate sources of information
- **compound class prior:** how likely the compound belongs to each of the three drug classes ("clinically important," "marginal," and "not clinically useful")
 - determined by decision makers
- **endpoint importance weights:** relative importance of each of the multiple endpoints on the decision
 - determined by decision makers

Drug Class Prior (1)

- Results of actual or **hypothetical** studies would be used for the compound
 - if there are multiple drugs available, then actual studies can be used to develop the priors
 - if there is limited information available, one could develop **hypothetical** scenarios of what a clinically important, a marginal, and a not clinically useful drug would look like
 - **developing such targets are a routine part of drug development in any case: this method requires that they be quantified**
- Conceptually, different decision makers could use different criteria
 - generally, however, they seem to come to a consensus on these different criteria

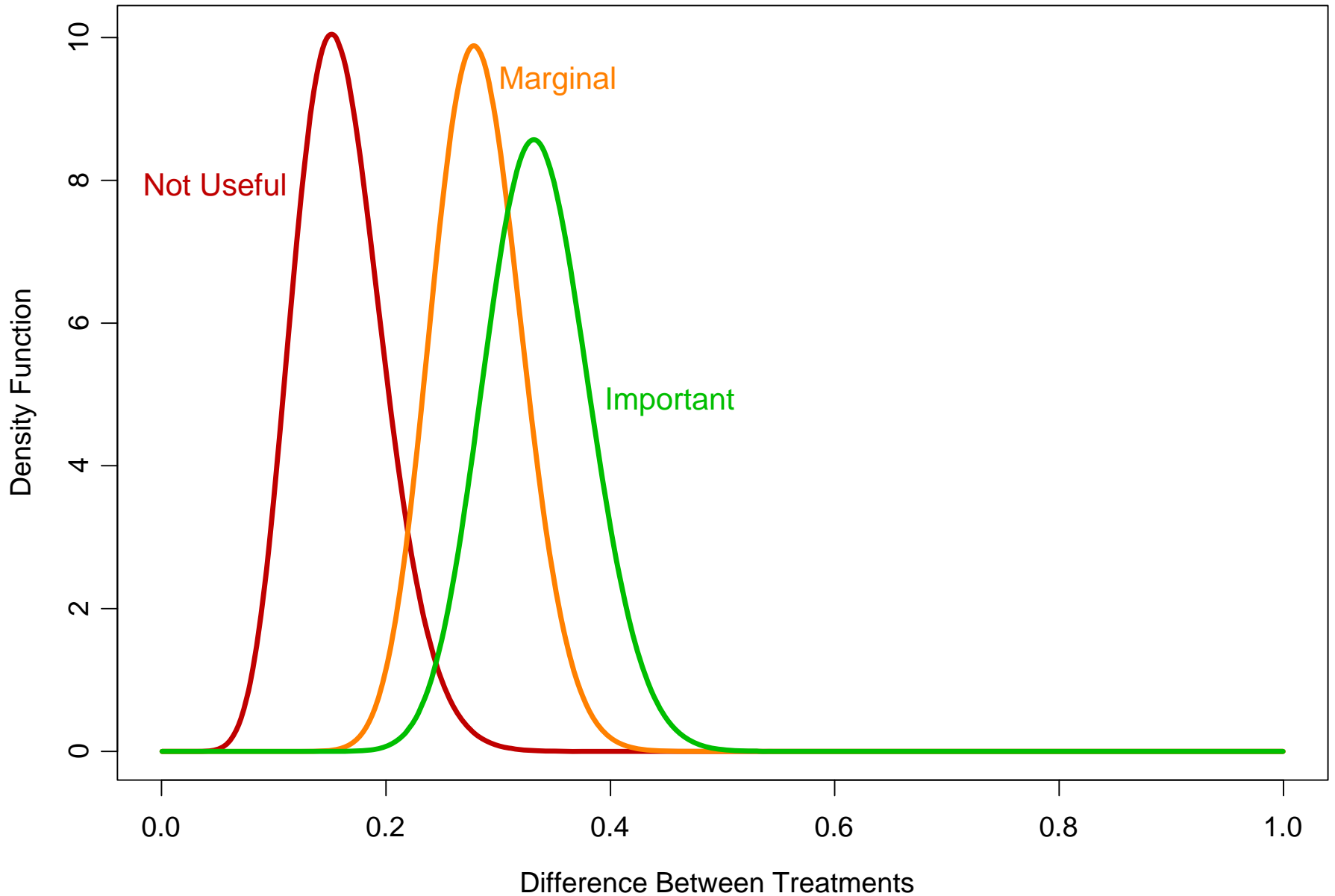
Drug Class Prior (2)

- **Certainly have a prior on efficacy**
- **Could also have criteria for other factors**
 - compliance in the clinical trial
 - side effects
 - fractures
 - patient reported outcomes

Drug Class Prior (3)

- **These priors could be based on**
 - a meta-analysis
 - multiple independent studies
 - a "representative" study
- **This is possibly the hardest information to obtain**
- **However, some criteria for what will lead to a "GO" or a "NO-GO" decision will be available in any case: this process forces it to be quantified**

Illustration: Drug Class Priors



Compound Class Prior (1)

- The **compound class prior** is how likely the decision maker(s) feel that the compound will be clinically important, marginal, or not clinically useful **before they have the data from the current study**
- It can be based on any approach the decision maker wants to user
 - can incorporate all sorts of "extra" information
- There is no right or wrong prior for this
- It can be different for different decision maker(s)
 - in fact, this was a feature that they seemed to like

Compound Class Prior (2): Example

	Clinically Important	Marginal	Not Clinically Useful
Optimist	75%	20%	5%
Typical	25%	50%	25%
Pessimist	1%	9%	90%
Non-informative (default)	33%	33%	33%

Calculation for a Single Endpoint

- To calculate the probability of belonging to a specific class:

$$\Pr(j) = \frac{w_j L(\theta_{jk}, \mathbf{x})}{\sum_j w_j L(\theta_{jk}, \mathbf{x})}$$

- w_j : compound class prior
 - $L(\theta_{jk}, \mathbf{x})$: likelihood of observing data \mathbf{x} , given the results of the k^{th} random draw of the parameters from the j^{th} drug class distribution, θ_{jk}
- Average over a large number of iterations, although for a specific single endpoint might have a closed form solution

Endpoint Importance Weights (1)

- **If there is more than one endpoint, then results can be presented both individually and as an overall synthesis across endpoints**
 - **The reality is that decisions are not going to be made based on a single endpoint**
- **This overall synthesis requires that there be some weighting of the different endpoints**
- **There is no right or wrong set of weights for this**
- **Different decision makers can have different endpoint weightings**

Endpoint Importance Weights (2): Example

	Spine BMD	Total Hip BMD	Side Effect Profile
Hip Focus	15%	60%	25%
Efficacy Focus	40%	40%	20%
Side Effect Focus	25%	25%	50%
Non-informative (default)	33%	33%	33%

Approach to Weighting of Multiple Endpoints (1)

Conceptually quite straightforward:

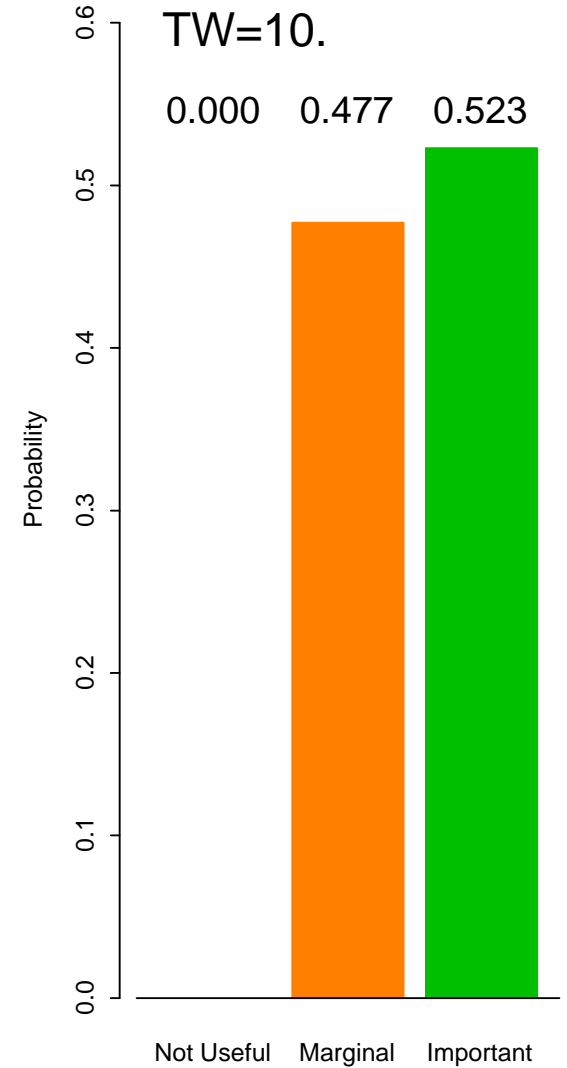
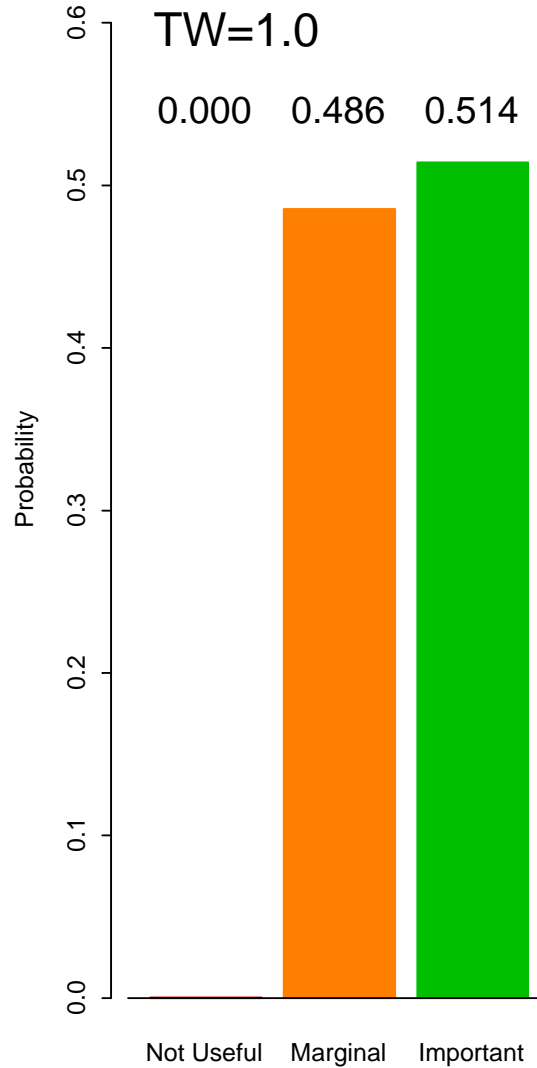
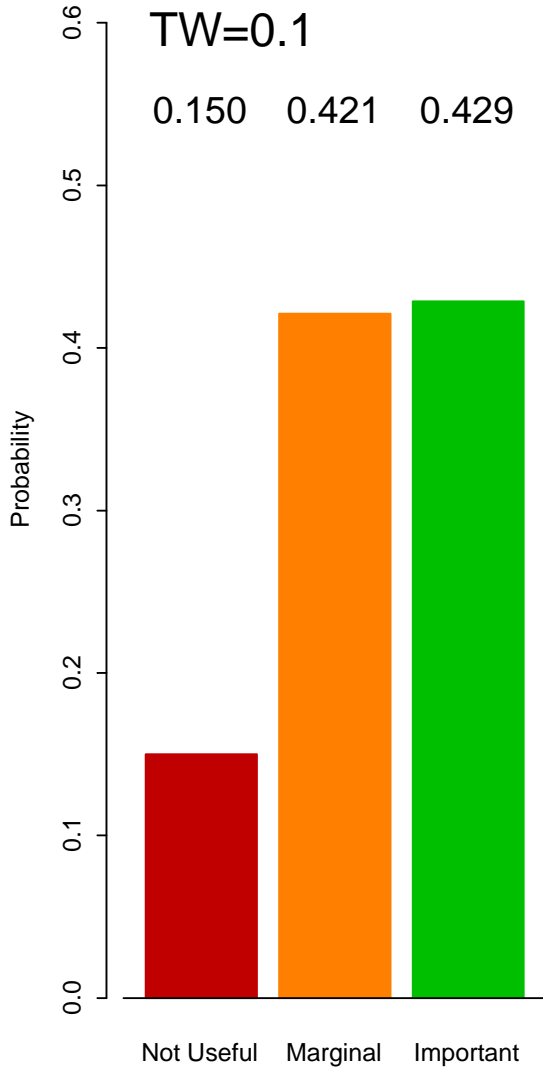
$$\Pr(j) = \frac{w_j \prod_i L(\theta_{ijk}, x_i)^{e_i}}{\sum_j w_j \prod_i L(\theta_{ijk}, x_i)^{e_i}}$$

- w_j : compound/class prior
 - e_i : endpoint importance weight for the i^{th} endpoint
 - $L(\theta_{ijk}, x_i)$: likelihood of observing x_i for the i^{th} endpoint, given the results of the k^{th} random draw of the parameters for the i^{th} endpoint from the j^{th} drug class distribution, θ_{ijk}
- Average over multiple iterations

Approach to Weighting of Multiple Endpoints (2)

- The total weight (sum of the e_i) used across all of the endpoints **does** affect the results
- Next slide shows impact of total weight (TW) with an extreme value, 0.1, to show convincingly that there is an effect, and more typical values

Effect of Total Weight Applied to Likelihood

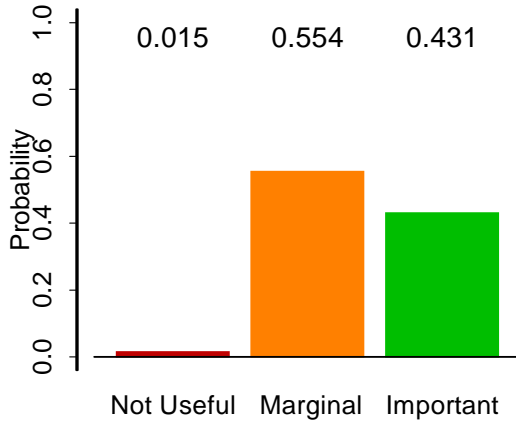


Approach to Weighting of Multiple Endpoints (3)

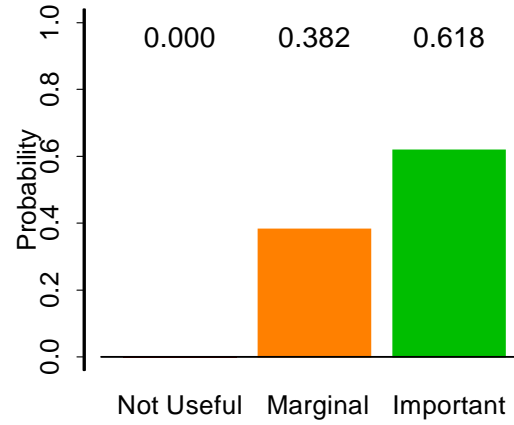
- Differences with reasonable values – 1 to 10, for example – are small and of no practical importance
- We set the total weights equal to the total number of endpoints included in the analysis, as this would be the total weight if all endpoints were counted equally without any formal weighting

Illustration of Results for Decision Maker

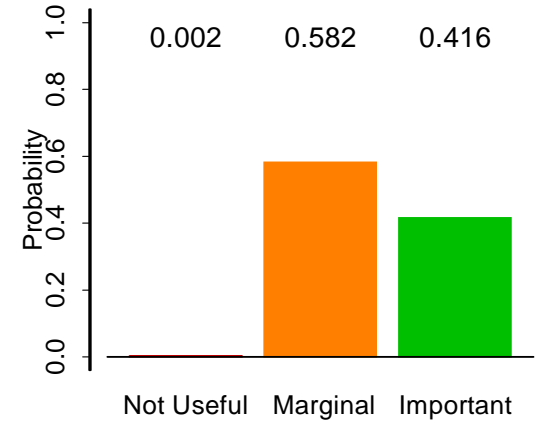
Endpoint 1



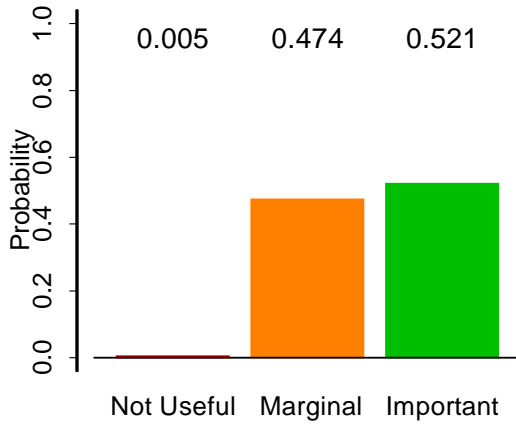
Endpoint 2



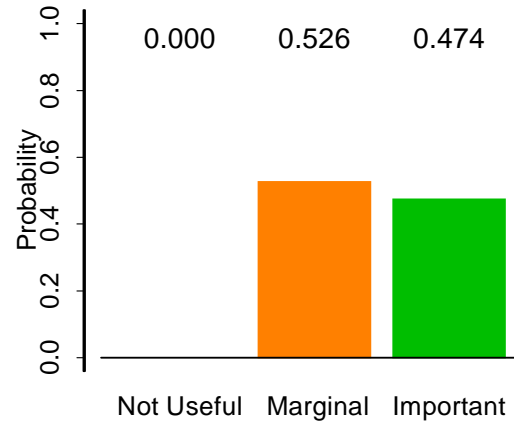
Endpoint 3



Endpoint 4



Weighted Likelihood



Some Obvious Extensions

- **Incorporating multiple drug class priors**
- **Incorporating individual weights and synthesizing across the decision makers**
- **Expanding to multiple stage decisions**

Summary (1)

- Presented an alternative approach to quantifying decision making for progressing a drug to the next stage of development
- Three types of information needed for model:
 - **drug class prior**: what "clinically important," "marginal," and "not clinically useful" means
 - **compound class prior**: how likely the compound belongs to each of the three drug classes ("clinically important," "marginal," and "not clinically useful")
 - **endpoint importance weights**: relative importance of each of the multiple endpoints on the decision

Summary (2)

- Much of this information is already available, especially the **drug class prior**
- Other information (**compound class prior; endpoint important weights**) may not be explicit, but is incorporated in the decision process in any case by decision makers: **this process makes such assumptions explicit**

Summary (3)

- **Critical statistical point** is that **the problem is changed** from one of predicting the probability of success to one of classifying how likely it is that this new compound belongs to one of three classes: a clinically important compound, a marginal compound, or a not clinically useful compound