



# Qualification of Biomarkers as Surrogate Endpoints of Chronic Disease Risk

Robert Temple, MD  
Associate Director for Medical Policy  
Center for Drug Evaluation and Research  
U.S. Food and Drug Administration

IOM

April 6, 2009



# Terminology

Even the name of the Committee illustrates the terminology problem: Qualification of Biomarkers as Surrogate Endpoints of Chronic Disease Risk. Well, let's separate two:

A surrogate endpoint is not intended primarily to measure risk; it is, once affected, intended to be an indicator of a clinical benefit (perhaps a reduced risk).

A biomarker can be enormously useful as a predictor of risk or predictor of responsiveness even if it cannot itself serve as a surrogate endpoint (e.g., many tumor surface markers) because it is not affected by the intervention.

Let me talk about surrogates first, biomarkers second (just a little).



# Definition of a Surrogate

Surrogate endpoint is not defined in the law or in any regulation. It is defined in the preamble to the Accelerated Approval Rule and mentioned under Fast Track in FDAMA (which more or less endorsed part of accelerated approval)

What is a surrogate?

- A surrogate endpoint of a clinical trial is a laboratory measurement or a physical sign used as a substitute for a clinically meaningful endpoint that measures directly how a patient feels, functions, or survives one. Changes induced by therapy on the surrogate endpoint are expected to produce changes in a clinically meaningful endpoint.
- The effect on the surrogate is, by itself, of no value to the patient. It is of value, i.e., it is a valid surrogate, only if the effect on the surrogate does lead to a clinical benefit.

The preamble points out that surrogates can have different degrees of support and that well-supported ones are currently a basis for approval.



# Why Do We Use Surrogates

Usually, we use a surrogate when the “real” endpoint is relatively rare or is very delayed, making it hard to study

- Great interest in CV medicine where endpoints are often delayed and, in any case, are rare (death, stroke or MI in survival in HT, elevated lipids, elevated BS; survival with CHF or abnormal rhythms). Trials in those settings take many thousands (Note, though, not many surrogates are used)
- Belief in “obviousness” of endpoint, even if survival gain hard to show, notably in oncology
- Other effective therapy can make survival effects hard to see (AIDS, many cancers) – again, leads to large studies.
- Existing effective therapy can make outcome trial hard; only NI study would be ethical, with all its problems

So, we use (or want to use) surrogates because

- They give a faster answer, sometimes related to urgent need
- Studies are smaller
- They’re more of a “sure thing” (you usually can anticipate result)
- We believe in them
- They may be possible when outcome trials are not



# What Kind of Surrogates Are There?

The list is endless and growing, but potential surrogates include

- Effect on “cause” – plausible to think endpoint is on causal pathway, but evidence can vary
  - Killing a microorganism – obviously should influence infection outcome
  - Reducing a well-documented risk factor: BP, LDL (or HDL) cholesterol, HbA1C – vary in “obviousness,” supported by epidemiology, animal models.
  - Shrinking tumors, decreased leukemic cells, i.e., getting rid of the abnormal growth leading to death
  - Decreasing VPB’s, the putative cause of fatal arrhythmias
  - Decreasing signs of inflammation (CRP, ESR) – controversial in heart disease
- Restoring the impaired function thought to cause bad outcome
  - EF in CHF
  - Coronary blood flow in AMI
  - Pulmonary function (Debate as to whether it is a surrogate)
  - Anatomical surrogates: infarct size, brain lesions
  - Normalizing an electrolyte abnormality

# Aren't There Some Other Endpoints That Are Problems?

When we consider problems of surrogates, keep several other endpoint categories in mind, because they can pose the same problems:

1. Not surrogates, but measurements or scales
  - Exercise tolerance, NYHA class
  - Depression, pain, dementia scales
  - Sleep lab results
  - Urine flow rates
  - FEV 1, other PFTs

These are not symptom measures like PROs, but formal scales that are thought to represent, or to correspond to, real clinical benefits.



# Other Endpoints with Problems

## 2. Non-Ultimate (Intermediate) Endpoint

- Symptoms of Hyperglycemia
- Progression to stage of AIDS, weight gain
- Surgical or medical relief of angina
- Prevention of MS relapses
- Relief of symptoms of CHF

In a sense, any symptomatic benefit is “non-ultimate;” they are real clinical benefits (decreased angina, improved exercise tolerance) but not the “ultimate” endpoints of improved survival or decreased irreversible morbidity.

Generally, the benefit would be valuable even without improved survival but not with decreased survival.

These have the same fundamental problem as surrogate endpoints: you don't know relation to ultimate endpoint and it could be adverse. 7

# What's the Surrogate Problem?

1. You only know what you measure

When you use a surrogate endpoint or an intermediate endpoint, there can always be a "surprise."

2. There are 3 main reasons for such surprises

1. The relationship between the surrogate and clinical event may not be causal, as supposed, but coincidental, or at least not directly related so that improvement of the surrogate does not influence the ultimate endpoint.
2. The surrogate or intermediate effect is only one measure of drug effect, the short-term effect presumed to be of value. Drugs may have other effects, not as obvious or prominent, and not as frequent, that may be adverse (or beneficial). This is not really a "surrogate" issue, but applies to any non-ultimate endpoint.
3. If there is both a favorable and adverse effect, whether there is net benefit or harm may depend on patient population or duration of use.



# Relationship Wrong

Whether incorrect relationships of surrogate to endpoint, as opposed to unexpected toxicity, have really been the major problem with surrogates, is not wholly clear, but there are some candidates:

Statins have far greater and earlier effects on recurrent MI and survival than is predicted by anatomic measures such as coronary artery cross-sectional area or plaque size; this may reflect greater complexity than the “cholesterol obstructs vessels” thought (my candidate is allowing vessel surface to heal and be less attractive to platelets).

torceptrapib had major HDL effect but worsened survival and CAD; suggests HDL alone not the whole story (maybe it has to be “packaged” correctly).

The whole heart failure story is sobering. There was a good deal of reason to think making the heart stronger would be good and an inotrope would have a benefit. In fact, many inotropes have been lethal and negative inotropes (BB's) have been helpful.



# Multiple Actions are Usual, Rarely Expected

- Multiple actions are common, not the rare exception, as we have discovered in almost every category of long term intervention in cardiology.
- We are usually surprised, because we've been focusing on the plausible (epidemiologically or pathophysiologically) beneficial effect of the drug throughout development and not other effects, because the adverse consequences of other effects are not easily predicted, and because we're optimists/believers.



# Multiple Actions

|                  | <b>Good</b>   | <b>Not Good</b>     |
|------------------|---------------|---------------------|
| Diuretic         | BP down       | K down; chol, BS up |
| Antiarrhythmics  | VPB down      | Mortality up        |
| Quinidine        | VPB down      | QT, TDP             |
|                  |               | Cy P450 II D 6 down |
|                  | NSR           | Mortality up        |
| Aspirin          | Thrombosis    | Bleeding up         |
| Inotropes, PDE I | down          | Mortality up        |
| Thrombolytics    | CI, ETT up    | Dissolve good clots |
| Cholesterol-     | Dissolve bad  | Rhabdomyolysis      |
| lowering         | clots         | Valvulopathy        |
| Fenfluramine     | CAD down      | CV events increased |
| COX-2 Sel        | Weight down   | (?)                 |
| NSAIDs           | Bleeding down |                     |



# Cardiovascular Interventions and Surprises

| Disease/Condition               | Surprise   |
|---------------------------------|--|
| Hypertension                    | Initial poor effect on coronary deaths, probably high dose diuretic toxicity   |
| Heart Failure                   | + Inotropes don't improve survival, may impair it -<br>Beta-agonists<br>Phosphodiesterase inhibitors<br>Digoxin - neutral effect<br>Dobutamine - repeat dose harmful<br>Flosequinan - harmful<br>Vesnarinone - dose dependent adverse survival<br>- Inotrope (beta-blockers) - improvement in survival |
| Post-infarct arrhythmias and SD | Increased mortality with several antiarrhythmics (except beta-blockers)  |
| CAD and cholesterol             | Pre-HMGC <sub>o</sub> A, no net effect on survival in most cases<br>MER-29 - cataracts<br>Late benefit - nicotinic acid  |
| Penetrating injury              | Transfusion with solutes may decrease survival   |
| Hormones post-menopause         | Adverse CV effects   |



# Subtler Differences

BP: A very good surrogate, and there is very strong evidence that all drugs that lower BP improve many hypertension outcomes, and we have just made final a labeling guidance that will give all antihypertensives outcome claims, but the drugs are not the same:

- Some (diuretics, BBs, ACEI s) treat CHF; CCB's don't
- Different effects on BS, lipids
- Some (AI I B's) slow rate of loss of renal function in type 2 DM; amlodipine didn't
- High dose diuretics may have less cardiac benefit because of hypokalemia

Not a failure of the surrogate, but drugs have multiple properties



# Unexpected Toxicity Is Not a Surrogate Issue

But it is particularly important when you've shown an effect on a surrogate. Unexpected toxicity could be important in many situations, including those where a drug had a clinical benefit. A reason for particular concern with surrogates is that without evidence of net benefit, the effect on the surrogate is of no value at all.

But

1. For a very well-established surrogate (BP), where benefit is assured, situation is very similar to situation where there is a demonstrated clinical benefit.
2. Some drugs provide benefits that are very modest, or can be obtained in many ways, so that their benefit would not override any important risk (antihistamines, simple analgesics).



## Attitude toward CV Surrogates

Concern over surrogates is of very long standing, especially in CV medicine, and very few are accepted:

- 1960's – “New York School” doubted benefit of BP lowering and asserted it would lead to strokes and heart attacks. VA studies “fixed” that, and many later studies added to comfort.
- Doubts about anti-arrhythmic benefits in 1980's led to CAST, showing increased post-infarction mortality with encainide, flecainide, and ethmosin. [Now, any anti-arrhythmic claim must include at least evidence of non-harm.]
- Drugs for CHF in 1980's (PDE inhibitors, beta agonists) showed worsened survival, so any symptomatic claim had to have evidence of benefit or non-harm. [We would not approve a symptomatic treatment without outcome data.] This means that, realistically, drugs for CHF all are studied in outcome trials.

## Attitude toward CV Surrogates (cont.)

- We have long approved drugs to lower LDL cholesterol but it was controversial until statins began to have outcome studies: 4S (1994), WOSCOPS (1998). In fact, a book was published just before those trials asserting it was all a sham. Results of ezetimibe trials have, to a degree, restarted the debate, but for statins, at least, evidence is overwhelming.

We still approve triglyceride lowering drugs but since torcetrapib, it's unlikely an HDL-lowering drug would be approved without outcome data. Reasonable support for LDL and triglycerides from trials of niacin, cholestyramine, fibrates, fish oil.

It can be debated, but putative "biomarkers" of benefit (IVUS, plaque measures, etc.) do not seem any better than the lipid measures (worse, I'd say), perhaps, because they don't reflect the state of the endothelium.

## Attitude toward CV Surrogates (cont.)

- We (and recent CR plus endocrine-metabolic AC) believe BS/HbA1C is a valid surrogate for BS control and microvascular benefit, but

NO ONE believes it is a validated surrogate of macrovascular benefit in type II DM

Controversy here is NOT the surrogate question but whether a chronically used drug in a fragile population needs outcome data to reassure us of lack of harm. Current interest in outcome data in drugs for:

- Diabetes (recent guidance)
- NSAIDs (current practice is large outcome studies)
- Probably, LDL lowering drugs (good-sized databases), although still approved based on surrogate
- Can probably think of others

## Attitude Toward CV Surrogates (cont.)

- A few more
  - K – we believe
  - Na – recent CRAC modestly supported improvement in marked hyponatremia
  - Venography
    - post-surgery
    - thromboembolism
    - prophylaxis
  - Creatinine (NOT really a surrogate, I think)
  - Toxicity surrogates
    - QT prolongation
    - Hy's Law
    - Uricosuric effect will cause ARF
  
- Finally, AMI
 

Apart from fatal AMI , hospitalization, pain, etc., an AMI has many of the characteristics of a surrogate endpoint (and we include asymptomatic ones).

Nonetheless, we are believers in the AMI endpoint (in MACE)



## Non-CV Surrogates

- BMD for osteoporosis prevention and treatment, but need fracture studies for fracture claims
- Testosterone suppression for prostate Ca Rx
- HIV viral load at 26/48 weeks for AI Ds drugs (Accelerated Approval at 26 weeks)
- Response rate in refractory cancer (Accelerated Approval). Moving toward TTP (still a surrogate)
- Endoscopic healing for duodenal ulcer



# I Illustrations

Antiarrhythmics

CHF

Anti-hypertensives

# Anti-Arrhythmics

CAST altered attitudes toward evaluation of anti-arrhythmics

VPB suppression had been basis for approval of anti-arrhythmics for symptomatic VPB's; CAST in post-MI people with  $> 10$  VPB/hour (marker of risk).

Post-CAST: Need to prove lack of adverse effect on survival. Document real effect on symptoms



# CAST

|                   | <u>Encainide/Flecainide</u> | <u>Placebo</u> |
|-------------------|-----------------------------|----------------|
| n                 | 755                         | 743            |
| Deaths and arrest | 63                          | 26             |

*Echt, et al, N Eng J Med 324:781-788 (1991)*



## Drugs for Heart Failure

Almost every expectation has been wrong

Inotropes (Beta agonists and PDEI s, dobutamine, flosequinan) harmful

Digoxin symptomatic benefit, no survival

Non-inotropes (ACEI 's, AI I B's) valuable

Beta blockers (contraindicated) turn out to be valuable

Diuretics probably valuable, but no real survival data, except with spironolactone on top of thiazide



# Promise

|                    | <u>Placebo</u>  | <u>Milrinone</u> |
|--------------------|---|------------------|
| n                  | 527   | 561              |
| Total mortality    | 127 (24%)<br>P = 0.038, nominal<br>P = 0.057, adjusted,<br>early stop | 168 (30%)        |
| Total CV mortality | 119<br>P = 0.016, nominal<br>P = 0.037, adjusted                      | 165              |

*Packer, et al, N Eng J Med 325: 1468-1475 (1991)*



# PROFILE

## Flosequinan vs Placebo in CHF

|           | <u>Flo 75</u> | <u>Plbo</u> | <u>Flo 100</u> | <u>Plbo</u> |
|-----------|---------------|-------------|----------------|-------------|
| Mortality | 40/206        | 43/238      | 201/964        | 138/937     |
| RR        | 1.05          | 1           | 1.48           |             |
| CI        | .68/1.62      |             | 1.19-1.84      |             |
| P-value   | 0.8254        |             | 0.0004         |             |

Flosequinan unequivocally improves exercise Capacity and CHF symptoms in NYHA Class II, III

Not tested in Class IV (can't exercise)



# Antihypertensives

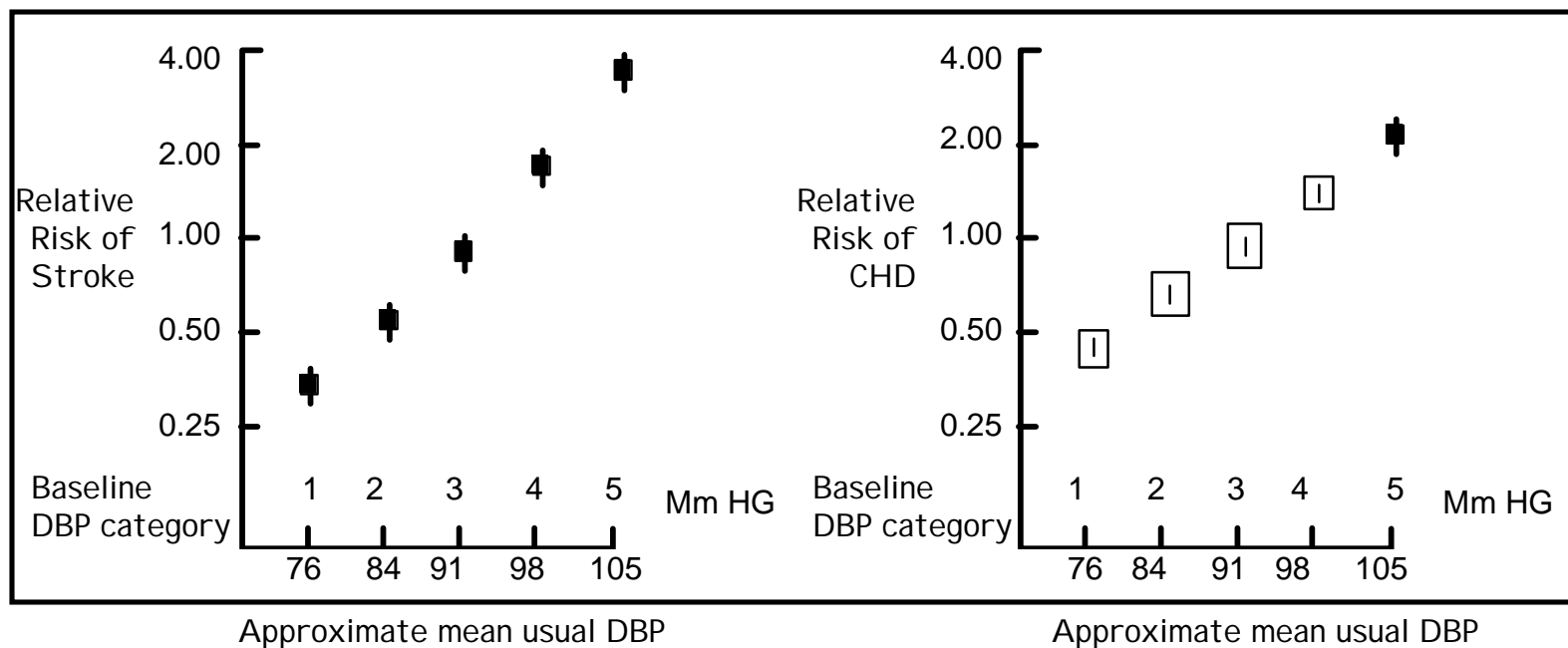
BP effect is a surrogate still, and enthusiastically, in use. We have proposed relabeling all antihypertensives to state that they all have effects on outcome (AMI, stroke, death).



## RELATIVE RISKS OF STROKE AND OF CORONARY HEART DISEASE, ESTIMATED FROM COMBINED RESULTS

Stroke and usual DBP  
(in 5 categories defined by baseline DBP)  
7 prospective observational studies: 843 events

Coronary Heart Disease and unusual DBP  
(in 5 categories defined by baseline DBP)  
9 prospective observational studies: 4856 events



Estimates the usual DBP in each baseline DBP category are taken from mean DBP values 4yr post-baseline in the Framingham study. See Table III. Solid squares represent disease risks in each category relative to risk in the whole study population; sizes of squares are proportional to number of events in each DBP category; and 95% CIs for estimates of relative risk are denoted by vertical lines.



## RESULTS IN ANTI HYPERTENSIVE TRIALS

Collins, et al  
Lancet 1990; 335:827-838

| <u>Endpoint</u>    | <u>Odds Ratio Reduction</u> |
|--------------------|-----------------------------|
| Stroke             | 42%                         |
| CHD                | 14%                         |
| Vascular Death     | 21%                         |
| Non-Vascular Death | 1%                          |

Best explanation for small effect on cardiac events  
is hypokalemia from high dose diuretics



## Surrogates: Yes/No

### PRO

#### A. Biologic Plausibility

1. Epidemiologic evidence consistent
2. Quantitative epidemiologic evidence (BP, lipids)
3. Animal model (BP, lipids)
4. Consistent with disease pathogenesis
5. Drug mechanism of action

#### B. Correlation of surrogate and clinical endpoint with another drug with:

1. Same pharmacologic effect (statins)
2. Different pharmacologic effect (antiHT)



## PRO (continued)

### C. Risk/Benefit

1. No alternative treatment for bad disease (Accelerated Approval)
2. Large amount of safety information (so unexpected effect unlikely), but there may need to be actual long-term controlled trial data
3. Short-term use
4. Difficulty of studying the desired endpoint (e.g., prevention of very late effects such as Alzheimer's dementia).



# CON

## A. Biologic Plausibility

1. A negative example (torceptrapib)
2. Inconsistent epidemiology

## B. Lack of any outcome data with the endpoint – Very hard case to make (except Accelerated Approval)

## C. Risk/Benefit

1. Alternative therapy known to affect survival (especially with different pharmacologic class)
2. Easy to study clinical endpoint (short-term outcome)
3. Long-term use in relatively healthy people
4. Low background event rate



## Special Concern

- Healthy population, prophylaxis (people not ill)
- Delayed impact, long-term use (maybe only choice)
- Relatively low event rate (limits benefit)
- Novel class of treatments (surprises likely)

All in all, surrogates are treated VERY cautiously

Question: Would a food-related claim have an easier time? Perhaps, but if the food or DS had to be taken in greater than past amounts, there could be concerns (beta-carotene in smokers; possible adverse effects of vitamins). We don't really know what daily broccoli, spinach, oatmeal in vast amounts do.



# Current State of Attitude toward CV Surrogates

All in all, there is a good deal of skepticism for surrogates without a great deal of data (BP, LDL, at least with statins). The anti-arrhythmic, heart failure, and torceptrapib experiences, all cases where the surrogate was VERY plausible, have had major effect.

## Other Surrogates

Very great interest in radiographic surrogates to predict neurological effects, tumor responses. In most cases the clinical studies are relatively easy to do. The tests will be where seeking:

- Prevention of very delayed neurologic effects
- Prevention of cancer

These involve very long trials, but also treatment of well people.



# Biomarkers - Definition

A broad term, without entirely-accepted definition. Generally, biomarker is used to refer to any physiologic, pathologic, or structural (micro or macro) property or activity that can be influenced by a drug, but that is only part of what the term means. It also refers to measurements that predict likelihood of an event (AMI, stroke, tumor recurrence), even if a drug does not influence them such as any risk factor, e.g., ECG finding, CRP levels, electrocardiographic findings, genomic pattern of tumor, proteomic or receptor patterns of a tumor.

Sometimes there is confusion as to whether the term refers to the measurement itself or to the change induced by a drug (i.e., a surrogate endpoint of a study).

In other words, is blood pressure the biomarker or a change in blood pressure the biomarker? I'll use biomarker to refer to the measurement itself and endpoint or effect to refer to the change in the marker that a drug might cause, prevent, or delay.



## Biomarker - Definition (cont.)

Thus, almost any measurement that is known or thought to be part of the pathophysiology of a disease and that could influence the nature of the disease or be affected by or influence the effect of a drug, can be considered a biomarker.

### 1. Cellular and circulatory receptors

This includes intracellular, membrane, or circulating receptors, whose levels can be measured (receptor could be enzyme, antigen, cytokine, hormone). Drug binding to these receptors can be of interest but levels of these could also predict outcome or response to treatment (e.g., tumor genetic or other markers)

## Biomarker - Definition (cont.)

### 2. Activity of an intrinsic molecule or externally introduced molecule

A next level biomarker is an enzyme, hormone, or cytokine whose activity can be measured and again, either predict outcome or be influenced by treatment (or both).

### 3. Level of an etiologic agent or anatomical features

The amount of a presumed etiologic agent, the most obvious being an infectious agent, or an anatomic abnormality (pathologic hallmarks of neurologic diseases, for example, or arteriosclerotic plaque structure or blood vessel characteristics (cross-sectional area, size). These too can be useful as measures of drug effect or as predictors of risk.

The importance and usefulness of such measures depends, of course, on how well understood the role of the biomarker is in disease and how well it corresponds to "later" (clinical) outcomes.



## Biomarker - Definition (cont.)

Finally, biomarkers include the still later measurements thought to be directly related to clinical outcomes, such as

Blood pressure

Total lipids, lipid fractions

Blood sugar, glycohemoglobin

Tumor size, tumor-specific antigens,

Coronary artery occlusion

These are often candidates for surrogate endpoints.



# Potential Role of Biomarkers, Other Than as “Pure” Surrogates

Broadly:

1. Efficiency and improvement in design of trials, identifying at risk populations, allowing preliminary screen for response, helping choose dose
2. Improved understanding of drug's effects and how to use it
  - Subgroup differences
  - Dose and dose-interval
  - Effects over time
  - Withdrawal effects
  - Pharmacodynamic interactions (favorable or unfavorable)



## Briefly, Biomarkers

Biomarkers that represent a drug's desired pharmacologic effect can contribute to "proof of principle," help choose dose, and encourage further studies, but I want to flag one particular potential, not as a surrogate endpoint for effectiveness, but as a way to identify people who should be studied, either because they will have the endpoints of interest or are capable of responding to the treatment, both methods of enrichment of the study population to increase study power.



# Selection of High Risk Patients

In outcome studies, the crucial determinant of study size is the number of endpoints the population will have. This is recognized and typically highest risk patients are studied early in lipid trials, BP trials (1967 VA study) and CHF.

CONSENSUS (first survival benefit in CHF studied NYHA Class III-IV patients. It needed only

253 patients

to show a survival benefit in < 6 months

All subsequent CHF studies had at least 2500 patients.



## Selection of High Risk Patients

We recognize risk stratification by LDL cholesterol, HDL cholesterol, BP, history of AMI, diabetes mellitus, and choose patients who will have higher risk, at least for initial studies.

But there are new “proteomic” measurements that seem to explain and amplify these predictors. CRP is a candidate, but here are others

Heeschen, et al. JAMA 2004; 291:435-441.

Examined ability of several blood factors to predict outcome (death + AMI) in population (placebo group in CAPTURE) who all had:

- Acute Coronary Syndrome
- >70% occlusion of at least 1 coronary
- Undergone angioplasty

I.e., they all look like similar high risk patients. But they're not.



## Selection of High Risk Patients

### b. Cardiovascular

Looked at predictive value of Placental Growth Factor (PIGF)

Soluble CD40 ligand (SCD40L)

Troponin

CRP



| Variable           | HR   | P     |
|--------------------|------|-------|
| Male               | 0.97 | .45   |
| Diabetes           | 1.24 | .62   |
| Smoker             | 0.67 | .23   |
| Hypertension       | 1.03 | .96   |
| CRP                | 0.98 | .94   |
| Troponin>0.01 mg/L | 1.83 | .03   |
| SCD40L >5mg/L      | 2.65 | .002  |
| PIGF >27 mg/L      | 3.03 | <.001 |

First 30 days risk of Death and AMI  
Cox proportional hazards  
Confirmed in 600 ER chest pain patients:  
PIGF>27 gave HR=4.80



## Selection of High Risk Patients

These are independent risk factors so that an ACS patient with all 3 predictors would have a 14.7 fold rate of events. (Note, CRP was fully accounted for by the other measures and so were other established risk predictors: diabetes, smoking, HT, maleness.)

The potential for doing a very small study in the high risk ACS population is fairly obvious. But it also reminds us that an apparently homogenous population can have very different people in it.

PIGF is a VEGF (vascular endothelial growth factor) and may be a factor in pathological angiogenesis; SCD40L is a measure of platelet activation; and troponin indicates myocardial damage, so their predictive value is not surprising



# Selection of Responders

Biomarkers can also be used to identify the subsets of a population capable of responding to a treatment. A classic case was the better response to BB's, ACEI's and ARBs of high-renin hypertensives. More recently we see

- Patients whose breast cancer has high levels of Her-2-neu receptor respond to trastuzumab
- Patients whose NSCLC has high levels of EGFR receptor respond to erlotinib.