# History of FDA Encouragement to Consider Subgroup Variability

Robert Temple, MD

JHU-CERSI Symposium

Heterogeneity

# History – A Little

There is long-standing recognition that responses to treatment (both effectiveness and safety) can differ among subgroups of the population, such as

- Demographic groups (age, gender, race)
- Disease severity
- Disease subgroups ( different causes)
- Disease duration
- PK/metabolic differences
- Concomitant illness
- Concomitant drugs

In many ways, there have been efforts to assess these in drug development, with initial focus on demographic difference.

# Heterogeneity and "Personalized Medicine"

Long known that not all patients respond identically to a treatment, but the reasons for this are becoming better understood. There are, of course, two quite distinct reasons for differences in response to treatment:

- Pharmacokinetic: people can differ in the rate at which they absorb or excrete the drug (kidney and/or liver function). Body size can also affect drug concentration. In addition, some people do not metabolize the active drug (giving higher blood levels) or do not form the active metabolite (clopidogrel), either because of genetic factors or because of concomitant therapy. Note that PK differences matter most when the drug dose is on the steep part of the D/R curve for S or E
- Pharmacodynamic: some people, because of genetic or pathophysiologic differences, respond to a treatment differently (high- renin vs low-renin hypertension response to ACEs, ARBs, BBS) EGFR positive vs negative NSCLC response to erlotinib

# Heterogeneity We Understand
# vs
# Looking for the Unexpected

## The world has changed dramatically with respect to PK heterogeneity:

- We now almost always know the major metabolic pathway for the active ingredient of a drug product and any active metabolite.

- We now almost always know how metabolism is affected by concomitant therapy and how the drug affects metabolism of other drugs [how fluoxetine would multiply tricyclic levels (desipramine, imipramine) because it is a CYP 450 2D6 inhibitor

- We also often (probably usually) have at least one blood level of the active ingredient for most patients [first suggested in the 1983 draft elderly guideline (final 1989), called a "PK screen"] and so can detect unexpectedly high or low blood levels and search for causes

# But Can Still Be Unexpected Subgroup Differences

Even if we now can anticipate/detect most differences in response related to PK, there are other potentially important differences that are not related to blood levels or metabolism, and in many cases we do not have any way to anticipate these [of course, when we can identify population subsets with and without a characteristic genetic or pathophysiologic marker at which the treatment is directed we can anticipate response differences. Indeed, it is usual to direct the treatment at the patient subgroup with a recognized response marker.]

But , for many diseases, we do not have the mechanistic understanding needed to identify the responders or to identify characteristics that will lead to adverse effects.

It therefore remains critical to examine population subgroups that MIGHT respond differently, and we continue to do that, focusing on Demographic Subsets, as the next slide shows. But there are other subgroups that could also have different responses and it is important to avoid unnecessary exclusion criteria (e.g., concomitant illness, concomitant treatment). In late 2013 FDA modified a MAPP directed at clinical reviews on monitoring INDs to emphasize the need to examine INDs for unnecessary exclusions (age, concomitant illness) that a study (Infusion study) of patients in NDAs from 2010 showed were common ( 71% excluded patients with a psychiatric disorder)

# Demographic and Other Subsets

History

| | |
|---|---|
| 1983 | Elderly Guideline draft |
| 1988 | Clin-Stat Guideline focus on subgroups |
| 1989 | Elderly Guideline |
| 1993 | Gender Guideline |
| 1993 | Do not start NDA review unless subset analyses done or readily available |
| 1994 | ICH Elderly Guideline; Q & A, 2010 |
| 1998 | Rule requires subset analyses (21 CFR 314.50), by age, gender and race in ISS, ISE |
| 2012 | FDASIA asks for report on whether demographic analyses are being done and communicated |

So we now always look, but there have not been very many M/F, old/young, and B/W difference (some, though).

# Subgroup Analyses

The literature is full of warnings about subset analyses and famous errors

- GISSI study of streptokinase showed effect only in patients with anterior MI. Later studies showed effect on MIs at all sites.
- ISIS2 showed beneficial effect of aspirin in patients born under all zodiacal signs except Libra and Gemini ( Peto's "watch out" example)

So… you must be very careful/cautious/skeptical about subgroup analyses

But, you still should look, and there have been critical findings.

# Pure PK

We routinely adjust doses for alterations in renal function (reducing dose in people with reduced function, but see NOACs later) and for drug-drug interactions (2D6 inhibitors increase tricyclic blood levels by 8-fold).

We generally ignore small blood level differences (but that is because most D/R and C/R relationships are relatively flat in the therapeutic range).

For toxic drugs (or maybe more broadly for drugs we know have steep D/R curves for effect and/or toxicity (cytotoxic oncology drugs) we often dose by weight.

# PK/PD

It is not always easy to know what effect small PK differences will have, e.g., we don't usually adjust dose for smaller size (gender), age-related falls in renal function, but in some cases small differences do matter.

- We know a variety of factors (genetic, food, drugs) affect warfarin blood levels, and we do worry about small changes; fortunately we have a good and easy way to measure the anti-coagulant effect, INR, so we monitor that.

- For hypnotics, we know size and renal function can affect blood levels (elderly tend to get higher levels and have often had lower recommended doses) although for Halcyon there was also a PD difference – the SAME blood levels caused more dizziness, etc. in older patients.

# PK/PD (cont)

For zolpidem, although we knew the same 10 mg dose would give higher average blood levels in women (smaller), there were no apparent differences in S or E, but a more sensitive PD marker led to lowering the dose.

We found that women were more likely than men to have morning blood levels high enough to impair driving performance, a realization made possible by a study that linked impairment on a driving test to blood levels over a certain threshold.

So, getting the right and very sensitive PD marker was crucial to realizing that the dose needed to be reduced in women.

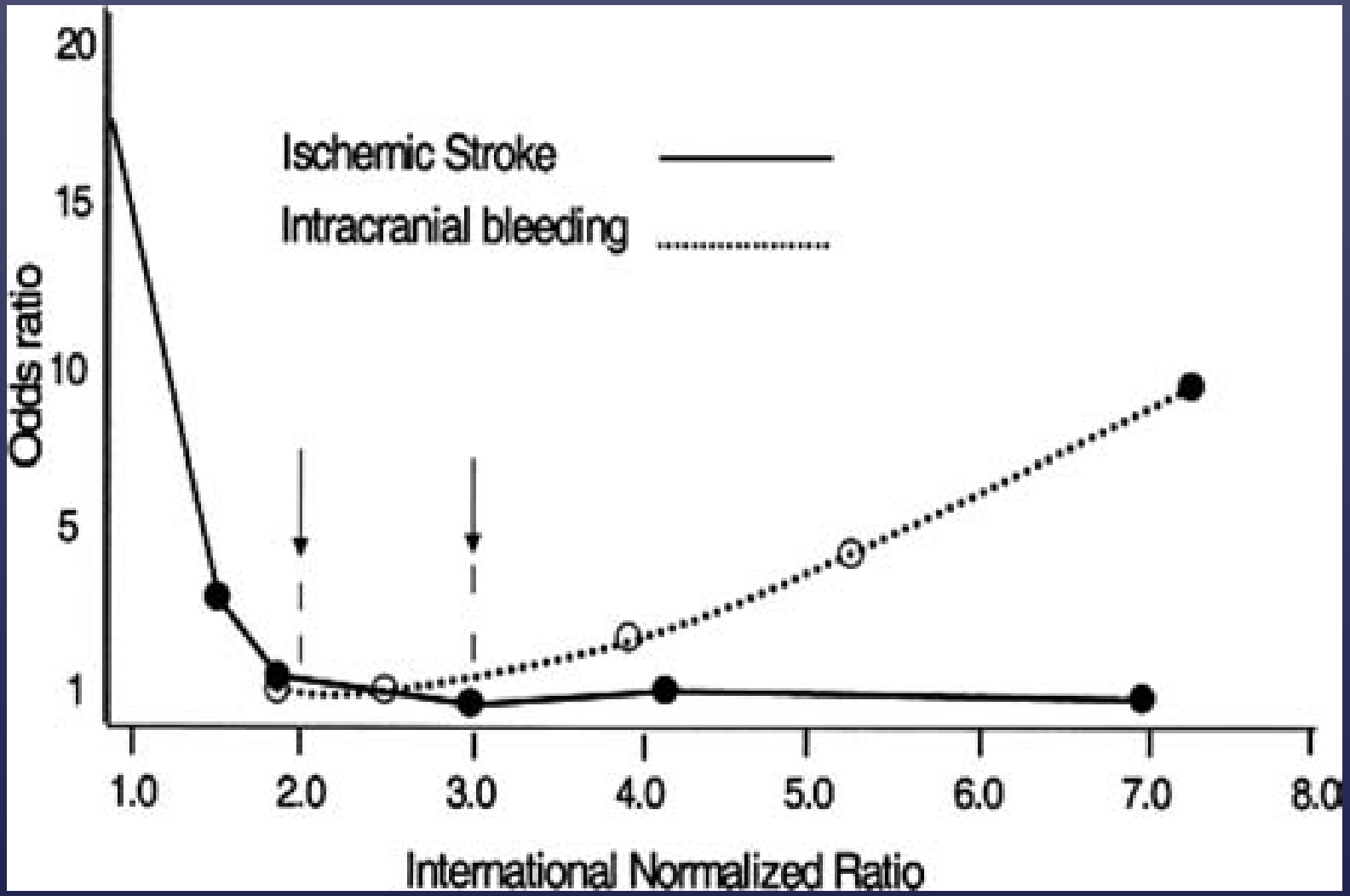
Amlodipine had far more cases of fluid retention in women given the 10 mg dose, an effect not seen in men, presumably reflecting greater blood levels, so the recommended dose was lower in woman

# PK/PD (cont)

For the NOACs, or at least dabigatran and edoxaban, we have extremely good data relating trough blood levels to the two critical endpoints when NOACs are used to treat AF

- thromboembolic stroke
- serious bleeding

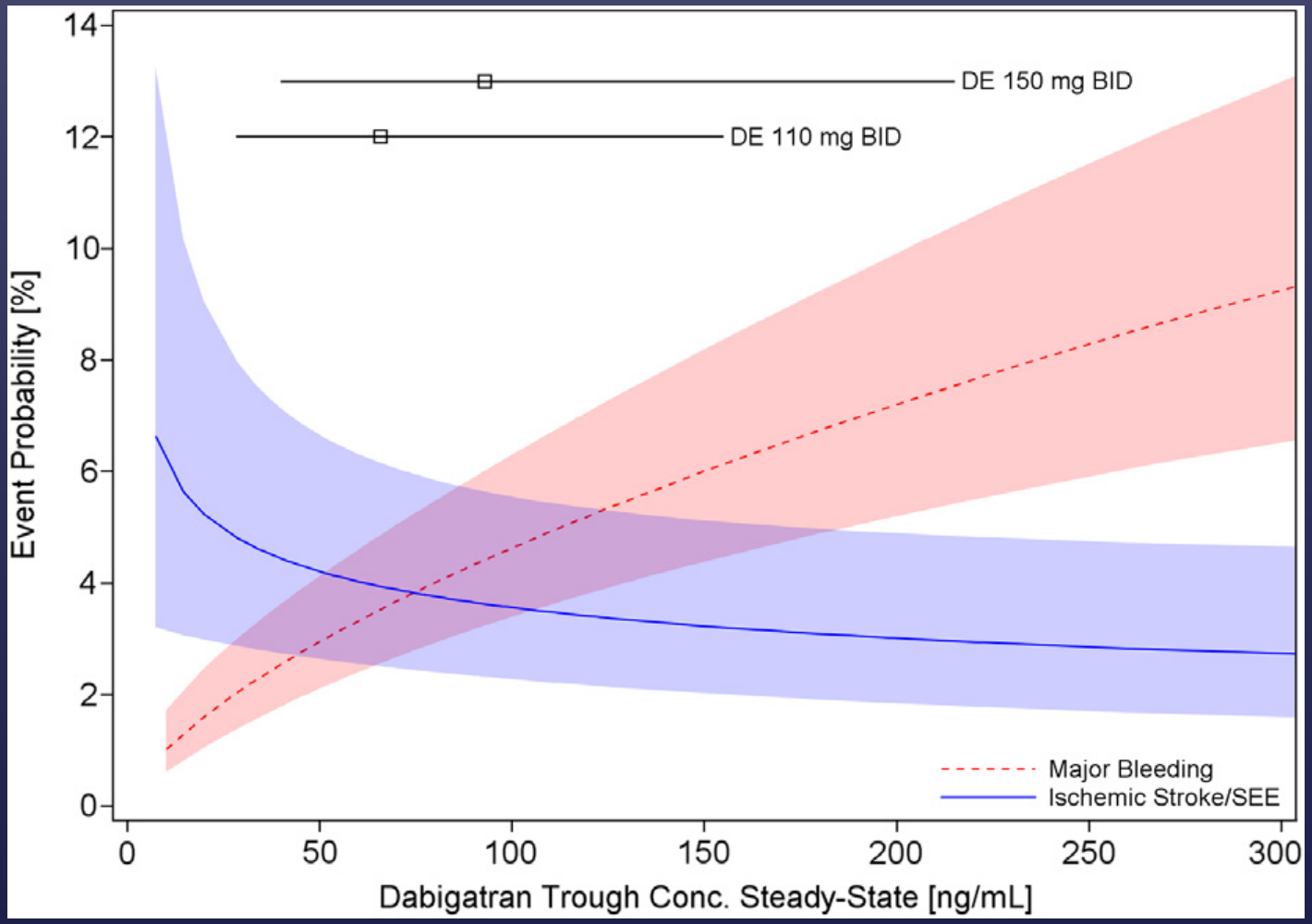
We knew that blood levels (translated into INR) corresponded to stroke rates and both intracranial and overall bleeding for warfarin. There is a “sweet spot” of INR 2-3 that optimizes stroke effect without too much bleeding. So the highly variable PK is really managed by assessing a relevant PD effect.


Ischemic Stroke
Intracranial bleeding
Odds ratio
20
15
10
5
1
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
International Normalized Ratio

# PK/PD (cont)

- NOACs (cont)

For dabigatran we saw, as for warfarin, that there was a threshold level for optimal stroke effect, about 75-150 ng/ml, with relatively little bleeding. This was clear from the clinical trial RE-LY, where the small difference between 150 mg and 110 mg had a marked effect (28% reduction) on stroke rate because the higher dose put almost everyone into the right concentration range. On the other hand, some people on 150 mg had blood levels greater than needed for optimal stroke reduction, at a cost of bleeding. And there is not yet any equivalent of an INR. It seems possible that measuring trough dabigatran levels could allow appropriate adjustment.


14
12
10
8
6
4
2
0
Event Probability [%]
DE 150 mg BID
DE 110 mg BID
Major Bleeding
Ischemic Stroke/SEE
0
50
100
150
200
250
300
Dabigatran Trough Conc. Steady-State [ng/mL]

# PD, But Not Fully Understood

The first rule of everything is

You're almost never quite smart enough

There are therefore cases of population differences where the differences surprise us. They are presumably PD-related, but not for reasons we know (at least not yet). A few illustrations:

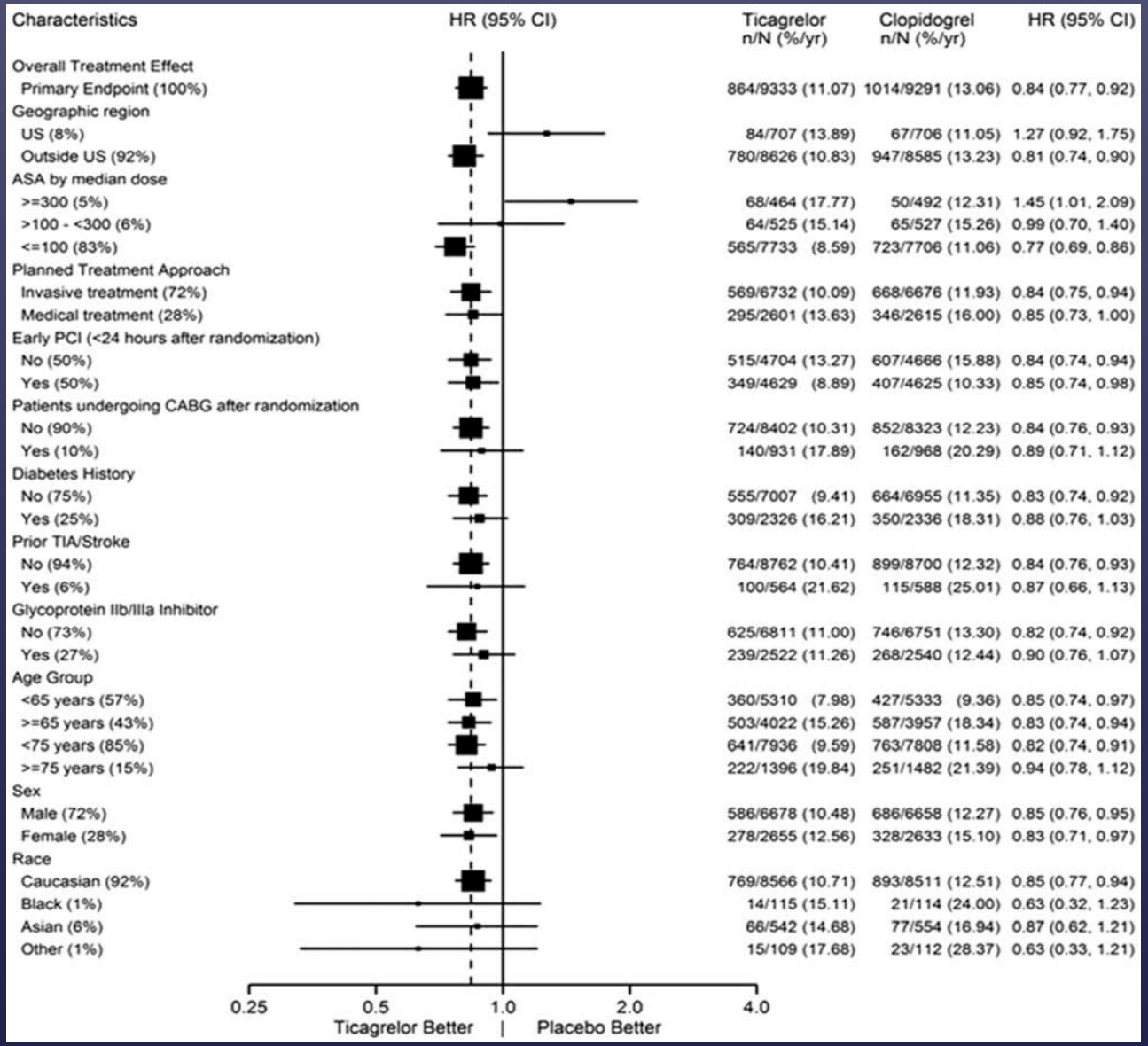
- Angioedema appears to be more common in blacks than whites but the increase in risk of angioedema from ACEIs is also greater in blacks.
- Alosetron, a drug for diarrhea – predominant IBS appeared to be effective only in women and was approved only for women.

# PD, But Not Fully Understood

- Ticagrelor

In the PLATO study of ticagrelor, early analyses showed an effect on CV mortality and non-fatal MI, in all regions but the US. This was shown to result from the use of higher aspirin doses in the US in about 50% of patients vs 7% elsewhere and ticagrelor's effect was reduced in people receiving higher dose aspirin. Corrected for ASA dose, results were similar in the US and elsewhere. The reason for the aspirin effect is not known, but the case illustrates how important it can be to examine subsets for possible differences.

I want to take this opportunity to suggest that Forest plots, shown for PLATO in the next slide, could be useful in settings beyond CV trials, where they are regularly used to examine subsets in a broad range of drugs for a wide range of subgroups.




**Source: Alison Blaus et al. Circulation. 2015;132:1425-1432**

# PD, But Not Fully Understood

- BiDil

Two early VA studies in CHF strongly suggested that there was a response (a strong one) to BiDil only in self-identified blacks.

There actually was reasonably persuasive evidence that the effect of BiDil in whites was small, at best. There were two previous studies, V-HeFT 1 and 2, that pretty convincingly showed, at best, a much smaller effect in whites.

BiDil

VHeFT 1

| | Overall (459) | | Blacks (128) | | Whites (324) | |
|---|---|---|---|---|---|---|
| | BiDil | Plbo | BiDil | Plbo | BiDil | Plbo |
| Annualized mortality | | | 9.7% | 17.3% | 16.9% | 18.8% |
| RR | 0.73 | | 0.34 | | 0.75 | |
| P | 0.09 | | 0.004 | | 0.11 | |

VHeFT 2

| | Overall (804) | | Blacks (215) | | Whites (574) | |
|---|---|---|---|---|---|---|
| | BiDil | Enal | BiDil | Enal | BiDil | Enal |
| Annualized mortality | | | 12.9% | 12.8% | 14.9% | 11.0% |
| RR | 1.23 | | 0.95 | | 1.48 | |
| P | 0.08 | | 0.83 | | 0.009 | |

# PD, But Not Fully Understood

We therefore allowed a trial in ONLY self-identified blacks, with quite spectacular results.

| | BiDil N=518 | Placebo N=532 | Risk Reduction |
|---|---|---|---|
| All Cause Mortality | 6.2% | 10.2% | 43% (p=0.012) |
| First CHF Hosp'n | 16.4% | 24.4% | 39% (p=0.001) |

As noted, we don't know why BiDil is more effective in blacks, but it clearly seems to be the case.

# Conclusion

There are enough cases of subset differences to suggest that it is always worth looking at subsets and worth including a broad range of patients

But, of course, be careful in conclusions (impact of zodiacal signs in ISIS I).

It is now SOP to do forest plots of any outcome study and they are often included in labeling. Our regulations call for analyses of effectiveness and safety results by age, sex, race, and other characteristics of interest. As noted earlier, Forest plots might be more broadly helpful tool