

# **Thrombolysis for Acute Ischemic Stroke?**

Jerome R Hoffman, MA MD

Professor of Medicine Emeritus  
UCLA School of Medicine

## **TPA for Stroke: Benefit vs Harm**

- **Does it benefit any patients?**
- **Does it harm any patients?**
- **Does it benefit more patients than it harms?**

## Defining Outcomes

- **Efficacy: Outcomes obtained under ideal circumstances**
- **Effectiveness: Outcomes obtained in “real world” circumstances**

## Threats to Efficacy- External Validity

- **Multiple studies:**
  - 1 (now 2) = “positive”
  - Most = neutral
  - A few = “negative”
- **How to explain ... ?**

# The Original RCTs

| RCT      | Date         | #   | Time | Mortality             | Fxn                         | Other                               |
|----------|--------------|-----|------|-----------------------|-----------------------------|-------------------------------------|
| tPA      |              |     |      |                       |                             |                                     |
| ECASS I  | Lancet 10/95 | 620 | 0-6h | 18 / 13%              | no difference               | (except ... "target population")    |
| NINDS    | NEJM 12/95   | 624 | 0-3h | 17 / 21%              | 13% better                  |                                     |
| ECASS II | Lancet 10/98 | 800 | 0-6h | 6.1 / 4.9%<br>= @ 90d | ≤minimal difference (p=0.3) | no difference in 20% treated in <3h |
| ATLANTIS | JAMA 12/99   | 613 | 3-5h | 11 / 7%               | no difference (1/3 each)    |                                     |
| PART A   | Stroke 4/00  | 142 | 0-6h | 18 / 4%<br>>22 / 7%   | no difference               | stopped early                       |
| STK      |              |     |      |                       |                             |                                     |
| MAST-I   | Lancet 12/95 | 622 | 0-6h | 27 / 12%              | no difference               |                                     |
| MAST-E   | NEJM 7/96    | 310 | 0-6h | 34 / 18%              | no difference               | 80% bad outcomes "killed off worst" |
| ASK      | JAMA 9/96    | 340 | 0-4h | RR ~2                 | much worse                  | 3 vs 4h                             |

# The Original RCTs

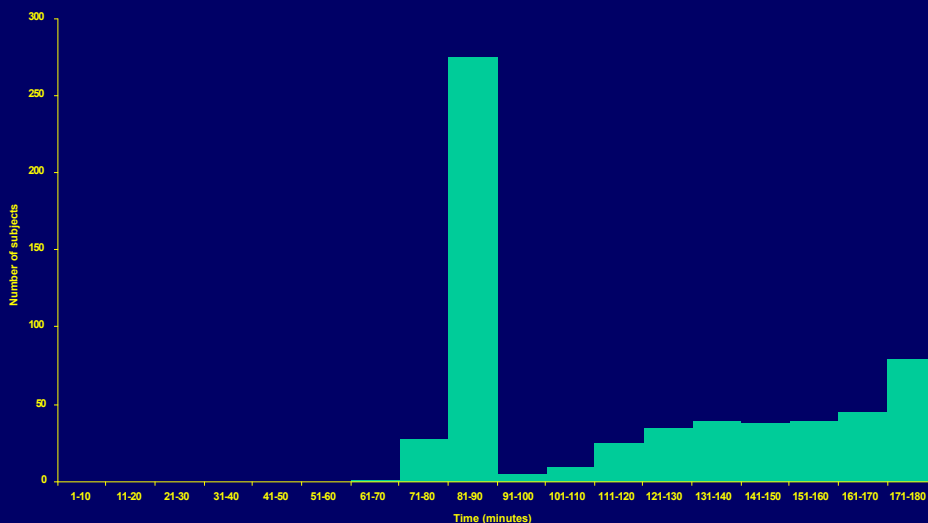
| RCT      | Date         | #   | Time | Mortality             | Fxn                         | Other                               |
|----------|--------------|-----|------|-----------------------|-----------------------------|-------------------------------------|
| tPA      |              |     |      |                       |                             |                                     |
| ECASS I  | Lancet 10/95 | 620 | 0-6h | 18 / 13%              | no difference               | (except ... "target population")    |
| NINDS    | NEJM 12/95   | 624 | 0-3h | 17 / 21%              | 13% better                  |                                     |
| ECASS II | Lancet 10/98 | 800 | 0-6h | 6.1 / 4.9%<br>= @ 90d | ≤minimal difference (p=0.3) | no difference in 20% treated in <3h |
| ATLANTIS | JAMA 12/99   | 613 | 3-5h | 11 / 7%               | no difference (1/3 each)    |                                     |
| PART A   | Stroke 4/00  | 142 | 0-6h | 18 / 4%<br>>22 / 7%   | no difference               | stopped early                       |
| STK      |              |     |      |                       |                             |                                     |
| MAST-I   | Lancet 12/95 | 622 | 0-6h | 27 / 12%              | no difference               |                                     |
| MAST-E   | NEJM 7/96    | 310 | 0-6h | 34 / 18%              | no difference               | 80% bad outcomes "killed off worst" |
| ASK      | JAMA 9/96    | 340 | 0-4h | RR ~2                 | much worse                  | 3 vs 4h                             |

## Threats to Efficacy- External Validity

- **Multiple examples of overly optimistic (single or grouped) small studies**
  - Magnesium
  - Ancrod
- **What does it mean when a couple of studies are “+” ... a bunch are neutral ... & a few are “-”**
  - Imagine flipping 1 penny ... or 10 different pennies

## Threats to Efficacy – Internal Validity

Artificial recruitment times in NINDS

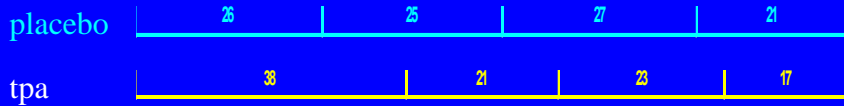


Marler, Neurology 12/00

## Threats to Efficacy – Internal Validity

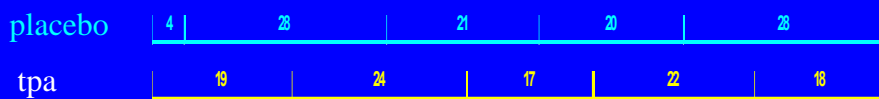
### Major differences in baseline severity

“TPA > placebo” in all NINDS patients at FINAL OUTCOME



In 91-180' subgroup, TPA >> placebo AT BASELINE

Least Severe Most severe



Marler, Neurology 12/00

## Threats to Effectiveness

- **Stroke mimics**
  - Common
  - Much more likely to “meet criteria” for tPA
- **Inaccurate CT reads**
  - Missed hemorrhage
  - Unreliable estimates of “large” territory
- **Pressure to enroll patients**

## Effectiveness Studies: “Great!”

- **Tiny ones**
  - “consistent with NINDS ... consistent with anything”
  - Industry sponsored
  - Methodologic concerns
  - Publication bias
- **3 big ones: STARS, CASES, SITS-MOST**
  - What they claim vs what they actually found

## Effectiveness Studies: “Awful!”

- **Non-selective, no publication bias, not ® sponsored**
- **Large #s, very frequent protocol violations**
- **Terrible outcomes (even when no protocol violation)**

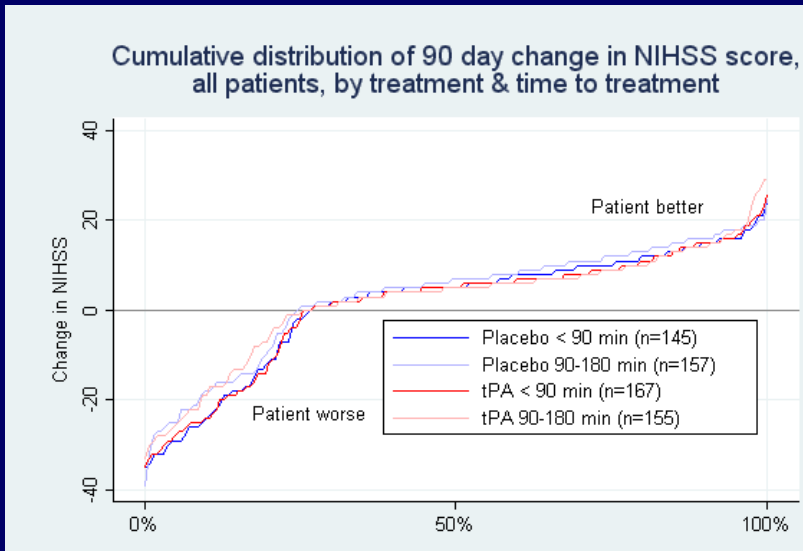
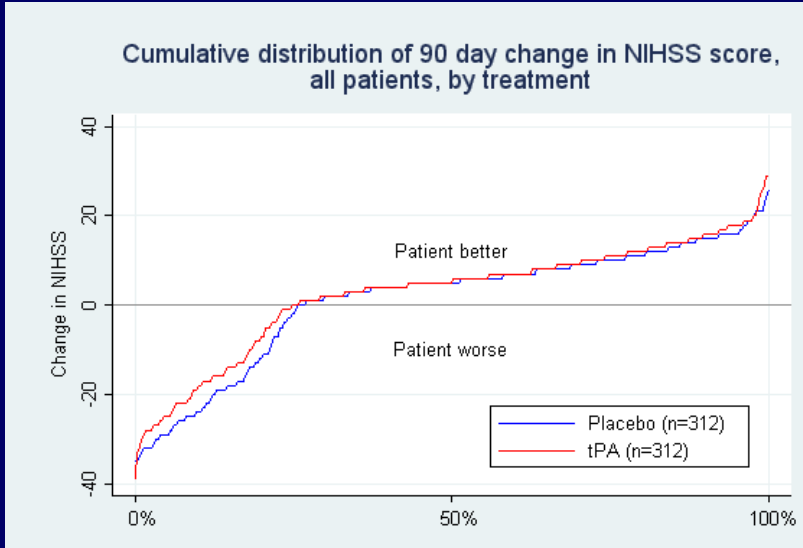
Katzan JAMA 3/00, Bravata Arch IM 9/02, Bateman Stroke 2/06,  
Deng Neuro 2/06, Dubinsky Neuro 6/06

## **The AHA Brouhaha**

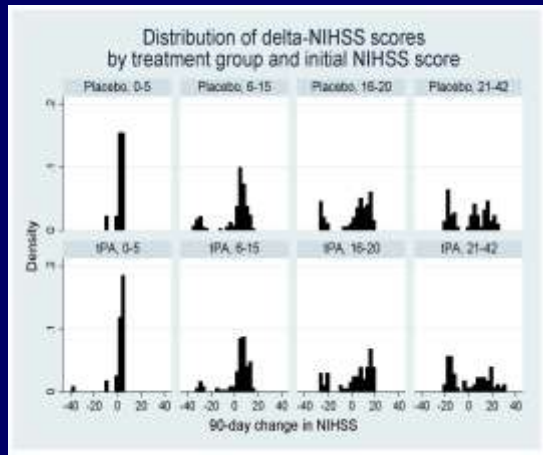
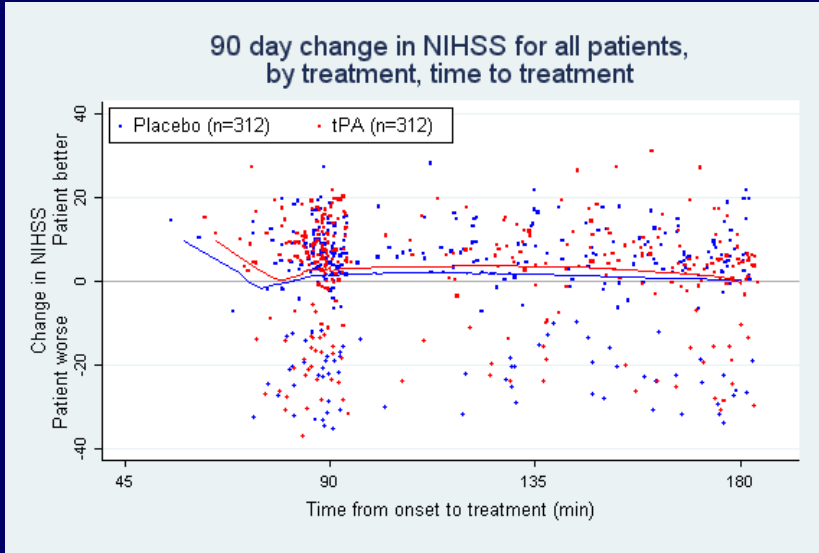
- **Money to the AHA**
- **Choice of experts**
- **“Disclosure” of conflicts of interest**
- **Suppression of dissent**

## **NINDS revisited**

After 10 years hidden ...  
let's look at the raw data







## Mean 90 day $\Delta$ in NIHSS score\*

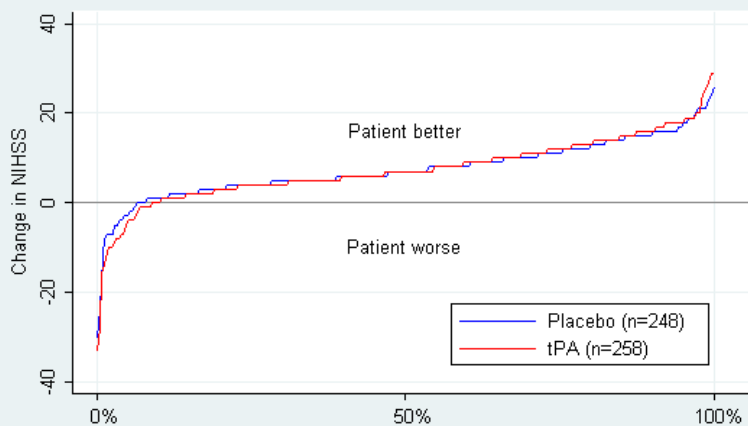
|              | Placebo | tPA | mean difference |
|--------------|---------|-----|-----------------|
| • $\leq 90'$ | 1.4     | 3.1 | 1.7 (-1.3, 4.8) |
| • 91-180'    | 1.1     | 2.1 | 1.0 (-1.9, 3.9) |
| • Total      | 1.2     | 1.4 | 1.2 (-0.7, 3.5) |

\*Positive values indicate improvement;  $\Delta$ -NIHSS ranges from -42 to +42

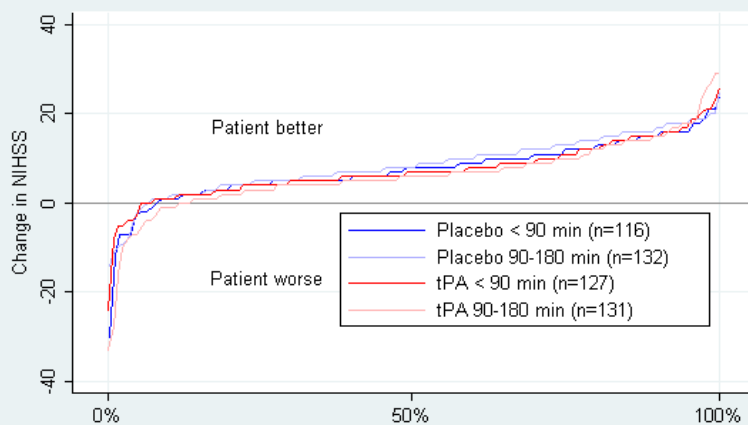
## % with $\geq 4$ point improvement at 90d

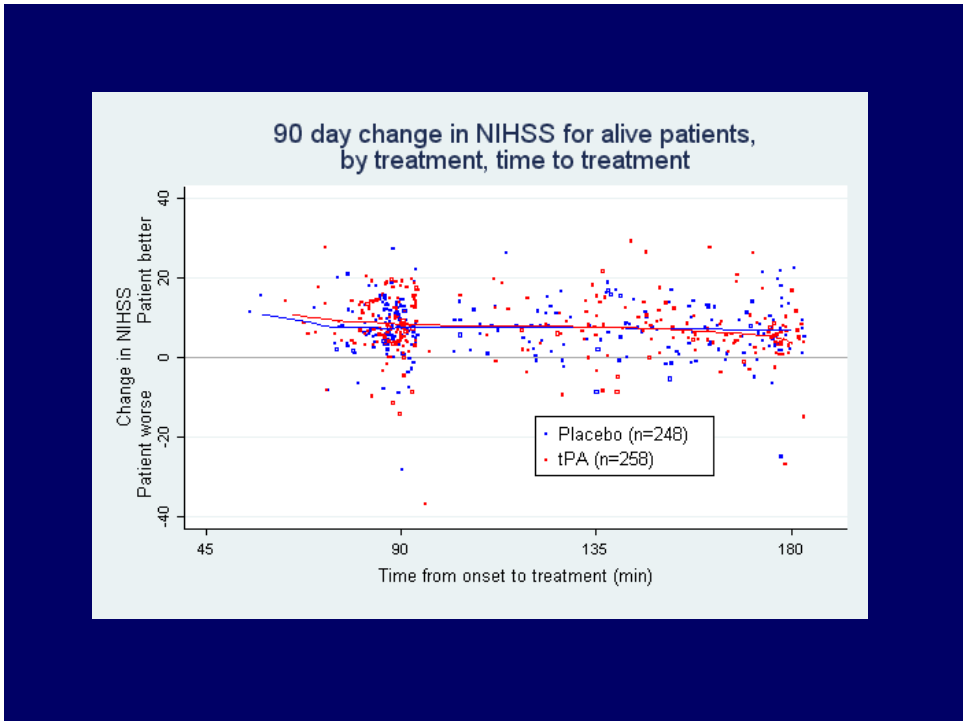
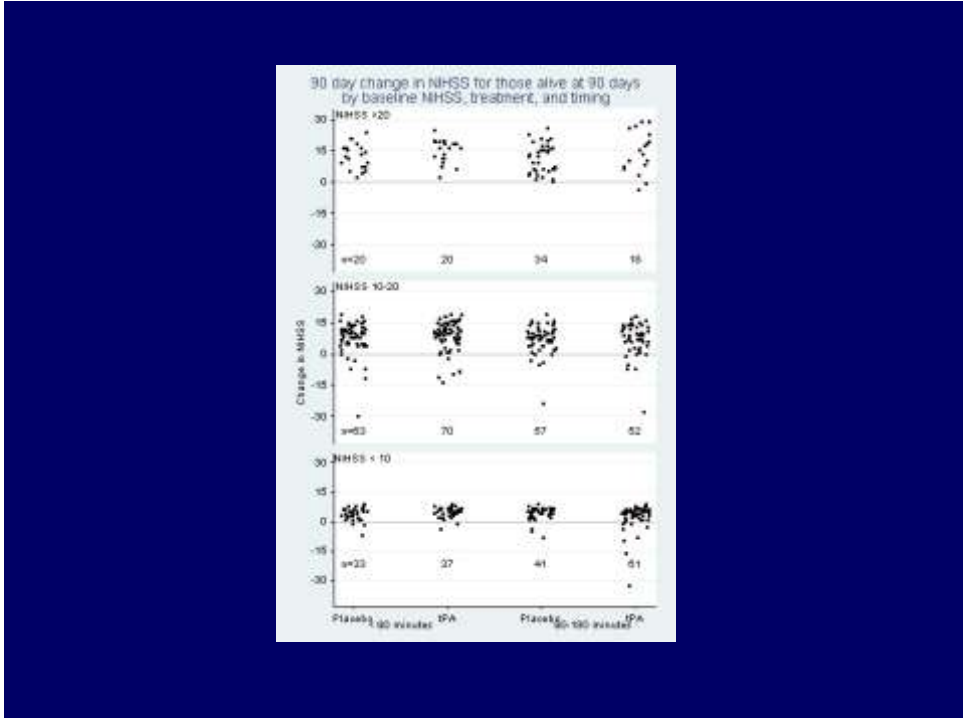
|              | Placebo | tPA | Difference (95% CI) |
|--------------|---------|-----|---------------------|
| • $\leq 90'$ | 64      | 66  | 2 (-9, 13)          |
| • 91-180'    | 62      | 61  | 0 (-11, 10)         |
| • Total      | 63      | 64  | 1 (-7, 9)           |

Cumulative distribution of 90 day change in NIHSS score, for those alive at 90 days, by treatment



Cumulative distribution of 90 day change in NIHSS score, for those alive at 90 days, by treatment & time to tx





## Mean $\Delta$ in NIHSS score in patients alive at 90d\*

|                      | Placebo (n=248) | tPA (n=258) | mean difference  |
|----------------------|-----------------|-------------|------------------|
| • $\leq 90'$ (n=243) | 7.4             | 8.5         | 1.1 (-0.7, 2.8)  |
| • 91-180' (n=263)    | 7.4             | 6.2         | -1.1 (-3.0, 0.7) |
| • Total              | 7.4             | 7.4         | 0.0 (-1.3, 1.2)  |

\*Positive values indicate improvement;  $\Delta$ -NIHSS ranges from -42 to +42

## % with $\geq 4$ point improvement among those alive at 90d

|              | Placebo | tPA | Difference (95% CI) |
|--------------|---------|-----|---------------------|
| • $\leq 90'$ | 80      | 82  | 2 (-8, 12)          |
| • 91-180'    | 78      | 73  | 6 (-16, 5)          |
| • Total      | 79      | 77  | -2 (-9, 5)          |

## Recent RCTs

| RCT       | Date              | #   | Time   | Mortality    | Fxn                   | Other             |
|-----------|-------------------|-----|--------|--------------|-----------------------|-------------------|
| tPA       |                   |     |        |              |                       |                   |
| ECASS III | NEJM 9/08         | 821 | 3-4.5h | 7.7 / 8.4%   | 3-6% absolute benefit |                   |
| Other     |                   |     |        |              |                       |                   |
| DPA       | Lancet Neuro 2/09 | 359 | 3-9h * | 6 / 11 & 21% | slightly worse        | * selected by MRI |

## IST-3 – The Largest RCT Ever

- n>3K, up to 6h p-sx onset
- 276 patients treated /evaluated 2-blind
- the other 90% done open-label, with “evaluation” by patient’s doctor ... or phone call (Q-with whom) ... or “letter gram” (Q-from whom) ... all without any effort at validation

## IST-3 – overall results

- No statistical difference in any of multiple outcomes @ 6mo
- However – a “positive trend” in neurologic function ... using a measure they’d explicitly rejected, in print, prior to finding the results
  - What is the clinical meaning of a “trend” in a study of >3K subjects?
  - How about for a 2<sup>o</sup> outcome ... done non-blind by a layman ... who is also non-blind to “was there any tx” ... using an unreliable measure the authors themselves rejected?

## IST-3 – results in 2-bl subgroup

- In the 276 patients treated /evaluated 2-blind ...
  - no hint of any benefit (no hint of any trend)
  - obvious major harms
    - mortality = 11 vs 7% (NNT<sub>K</sub>=25)
    - sx-ICH in 7 vs 1%
    - fatal ICH in 4 vs <1%
    - increase in non-ICH swelling, deterioration *not* due to bleed or swelling (10 vs 8%), & need for ICU (24 vs 17%)

## IST-3

- Mortality at 6mo = 27% in both groups
  - suggesting that even though TPA killed a few early, it is the *underlying stroke severity* that dominates ultimate outcome

## IST-3 – Authors Conclusions

“Offer tPA to all patients up to 6 hours!”

Laudatory editorial agrees!



# Clot Removal

## Clot Removal - Multi-MERCI

- 164 (non-consecutive) pts
  - Cherry-picked
  - Excluded “couldn’t reach the clot” from Results
- NIHSS = 19 (IQR 15-23)
- $\leq 8$ h p-sx onset
- “successful recanalization” in 57% ... 69% p-adjunctive IA tPA

Smith, Stroke 4/08

## Multi-MERCI

- symptomatic ICH in “10%” [16/13=12%]
- “clinically significant procedural complication” in “5.5%” [9/131=7%]
- 90d mRS = 0-2 in 36% [50/39% in NINDS]
- 90d mortality = 34%

Smith, Stroke 4/08

## Recent “Clot Removal” studies\*

- 3 RCTs vs various control groups
- Just over 1K subjects
- All in NEJM 3/13
- None showed *any* benefit (with trends toward worse outcomes)

## Clot Removal – Kidwell NEJM 3/13

- N = 118, <8h p-sx (mean 5.5h) randomized to clot removal or “standard care” ▲
- overall 90d mortality = 21%
- symptomatic-ICH in 4%

## Clot Removal – Kidwell NEJM 3/13

- 58% c “favorable pattern” on MRI: “small infarct core/ lots of salvageable tissue”
- “successful recanalization” c embolectomy in 2/3, independent of y/n “favorable” MRI
- “favorable” pattern pts did better (by some measures)
  - this was *equally true in controls* as in intervention group

## Clot Removal – Kidwell NEJM 3/13

- no difference in MR-reperfusion @d7, by tx group (50% each) or by y/n “favorable” pattern (ditto);
- no difference by tx in *90d mean mRS* (3.9 each), or in subgroups c or s “favorable” pattern (3.9 vs 3.4 & 4.0 vs 4.4) [so “favorable” pts did *better* without embolectomy!];
- no difference in % c *90d mRS*  $\leq 2$  (19% each), + no difference by y/n “favorable” (either between or within tx groups).

## “Clot Removal” studies – Conclusions of Authors

“These first-generation devices aren’t good enough. Just wait until we use the newer ones!”

## A Few Sources of Bias\*

- Methodologic
    - **Unreliable outcome measures**
      - mRS “by letter-gram” or phone
      - “stage shift”
    - Changing methods / changing outcomes
    - Selective rejection of methods ... depending on results
- \**Bias* in its standard English-language usage

## A Few Sources of Bias\*

- Methodologic
    - **Selective metaanalyses**
    - **Metaanalyses of outcomes not planned in the original studies**
- \**Bias* in its standard English-language usage

## A Few Sources of Bias\*

- Interpretive
  - **Publication bias**
    - Of studies
    - Of editorials
    - Of commentary within studies
  - **Selective rejection of studies** ... if their results are negative (b/c they're negative!)
    - \**Bias* in its standard English-language usage

## A Few Sources of Bias\*

- Interpretive
  - Conflating theory with evidence (DOOs with POOs)
    - MRI-based patient selection
    - Association vs cause
    - Ecological fallacy – the *group* does better ... but *not* the patients of interest (those who got tPA) within the group!
  - Changing the theory when it doesn't work
    - “Clot-busting” ... but no effect in 1<sup>st</sup> 24h, so ...
    - \**Bias* in its standard English-language usage

## A Few Sources of Bias\*

- Interpretive
  - Illogical “proofs”
    - “every 15’ helps” (Saver, JAMA 6/13)
      - What they found
      - What one would expect with a *completely worthless* treatment
    - “the NNT is really 3 ... or 2 ... or less than 1!”
  - **Everything is “positive” ... regardless of results**
    - \**Bias* in its standard English-language usage

## A Few Sources of Bias\*

- **Ethical**
  - Subtly changing the outcome measure ... and pretending you didn’t
  - Excluding protocol violations in an effectiveness study
  - Including only some of the participants
  - Changing methods & outcomes mid-stream ... once things look bad
  - “It’s unethical to do another RCT of <3h ... but fine to keep trying at 6h”
    - \**Bias* in its standard English-language usage

## **TPA for Stroke: Benefit vs Harm**

- **Does it benefit any patients?**
- **Does it harm any patients?**
- **Does it benefit more patients than it harms?**

## **TPA for Stroke: Benefit vs Harm**

- **Does it benefit more patients than it harms?**
  - Overall?
  - In *very early* patients?
  - With specific CT/MR findings?
  - Via IA route?
  - In expert hands?

**DO WE KNOW?**



## TPA for Stroke

➤ **Why are there no more 0-3h RCTs?**

“We do not know how another trial would turn out, and if we do not come out ahead, we would have a terribly self-inflicted wound....[Another study] may be a good thing for America, but it wouldn't be a good thing for us.”

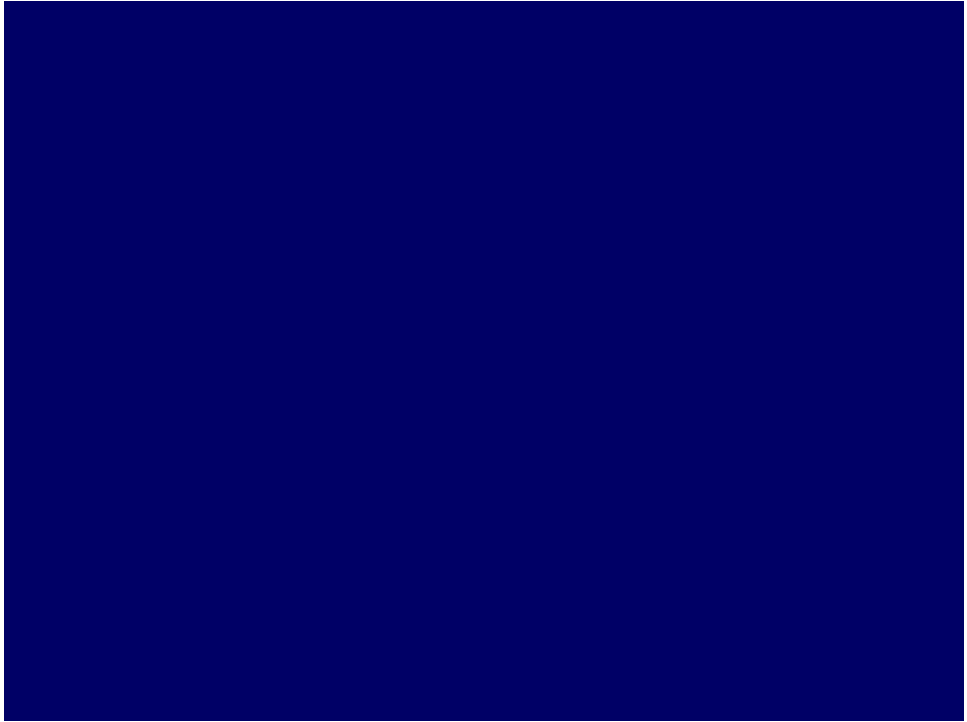
*Elliot Grosbard, Genentech scientist*

➤ **Why are there more >3h RCTs?**

## When to Use an Unproven Therapy

- Current outcomes are universally terrible
- Little possible harm from proposed intervention
- Potential for enormous benefit
- Likelihood that effectiveness = efficacy

**WHAT ABOUT 0 OUT OF 4?**



## Cleveland QI: “as good as NINDS”

- **At 9 Cleveland Clinic hospitals only**
  - >1/4 @ 1 UH
- **47 patients total (~2.5%)**
- **No chart review Methods**
- **Protocol violations in “only” 19% (8-30%)**
- **Nothing about *outcomes!***

Katzan, et al. Stroke. 2003 Mar; 34:799-800

## Threats to Efficacy: Other Studies

- **The STK studies**
  - Are they relevant?
- **ECASS and “target population” data**
  - Stands the scientific method on its head
  - Disproportionate “adjustment” in tPA group with bad outcomes
  - Not a criterion in NINDS
  - *If it were important ... impossible to ID on CT!*

## Effectiveness: STARS

- Large, with good results
- **NOT** an efficacy study: all = ATLANTIS sites
- **58** of 83 sites - what about the other 25?

## Effectiveness: CASES

- ®-sponsored registry, with “investigators” paid for cases voluntarily submitted
- “... estimate = 84%” of patients who got tPA
- No 90d f/u on 13%, so outcome “imputed” from time of hospital discharge
- No independent measure of actual outcomes

## Effectiveness: CASES

- **Abstract:** “excellent clinical outcome in 37%”
- **Methods:** “excellent” defined as mRS = 0-1
- **Results:** “... not significantly lower than NINDS (36.8% vs 39.9%), & similar to other series”
- **Interpretation:** “36.8%”

## CASES: “Adjusted mRS”

- “...difference between baseline and 90d mRS ...  
for example, a patient with baseline and 90d mRS = 3  
would be rated ‘excellent.’ ”
- **Abstract:** Not mentioned
  - **Methods:** Not defined or mentioned
  - **Results:** Not mentioned (except in Figure)
  - **Interpretation:** Not mentioned
- 
- **NINDS (tPA or placebo comparator):** No such measure

## CASES: mRS 0-1

- **The actual “primary outcome”**
- **Achieved in 31.8%**
  - **not mentioned in abstract, results, or interpretation**
  - **most consistent with placebo in NINDS**
  - **< achieved in 40K placebo patients in CAST / IST**
- **Add in selection bias, measurement bias, loss to f/u ...<30%? ... <25%? ... <20%? ... ???**

## Effectiveness: CASES

- **Abstract: “excellent clinical outcome in 37%”**
- **Methods: “excellent” defined as mRS = 0-1**
- **Results: “... not significantly lower than NINDS (36.8% vs 39.9%), & similar to other series”**
- **Interpretation: “36.8%”**

