

“10% of the time, it works every time”

## Recognizing sloppy science

Jim Broach  
Penn State College of Medicine  
Reprise of talk given by C. Glenn Begley,  
former Director of Oncology, Amgen

Between 2002-2012, Amgen was not able to reproduce 47 of 53 seminal publications. These were publications that reported something completely “new”

In the majority, data was not reproduced by the original investigators with their reagents in their lab

Amgen’s experience is not unique....

# Begley's position statement

- These results do not challenge the validity or legitimacy of the scientific method
- Not talking about fraud: the subject is scientific-laziness, sloppiness, ignorance, exaggeration, desperation
- The vast majority of investigators want to do the right thing
- That this debate is occurring in public confirms the strength our scientific system

**BE SKEPTICAL**

# High-profile studies typically fail at multiple levels:

## Begley's six criteria for judging scientific reports:

### 1) Were studies blinded?

Almost never

### 2) Were all results shown?

Typically not      “representative examples” & data selection bias

western blots that show only a slice; no size markers

### 3) Were experiments repeated?

Often not      westerns/immuno-precipitation usually only performed once  
typically only use 1/2 siRNAs and in 1/2 cell lines  
confusion between replicates and independent experiments

### 4) Were positive and negative controls shown?

Typically not

### 5) Were reagents validated?

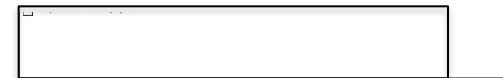
Frequently not      IHC with a polyclonal anti-peptide Ab  
small molecule inhibitors

### 6) Was the analysis appropriate (e.g. cell growth/statistical tests)?

Typically not

# Were all the results shown?

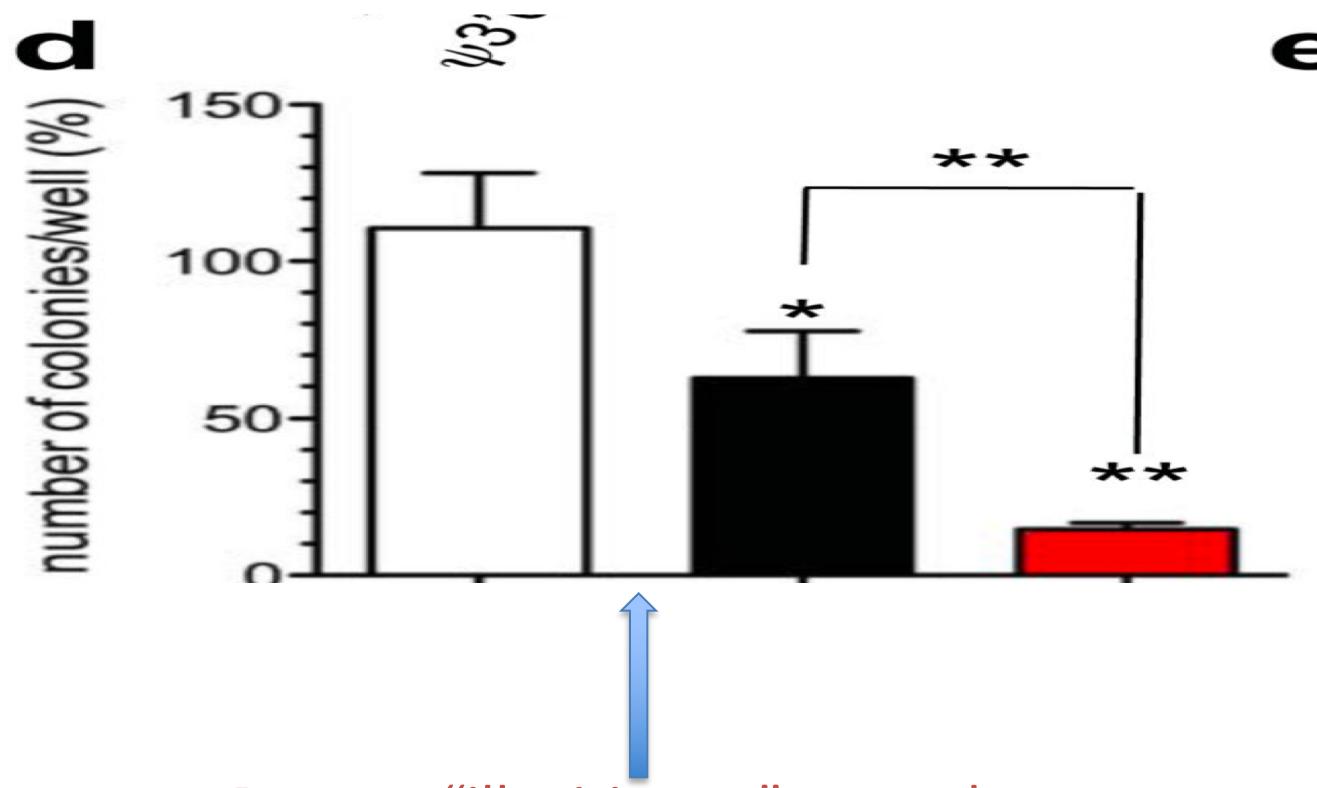
Versus



Investigators hide  
poor experiments  
by failing to show all  
the data

“Anti-EPOR” antibodies

# Were positive and negative controls shown? Beware using polyclonal anti-peptide antibodies



Beware “illegitimate” controls  
Here competition experiments are not the appropriate control  
- an apparent control may not be a control

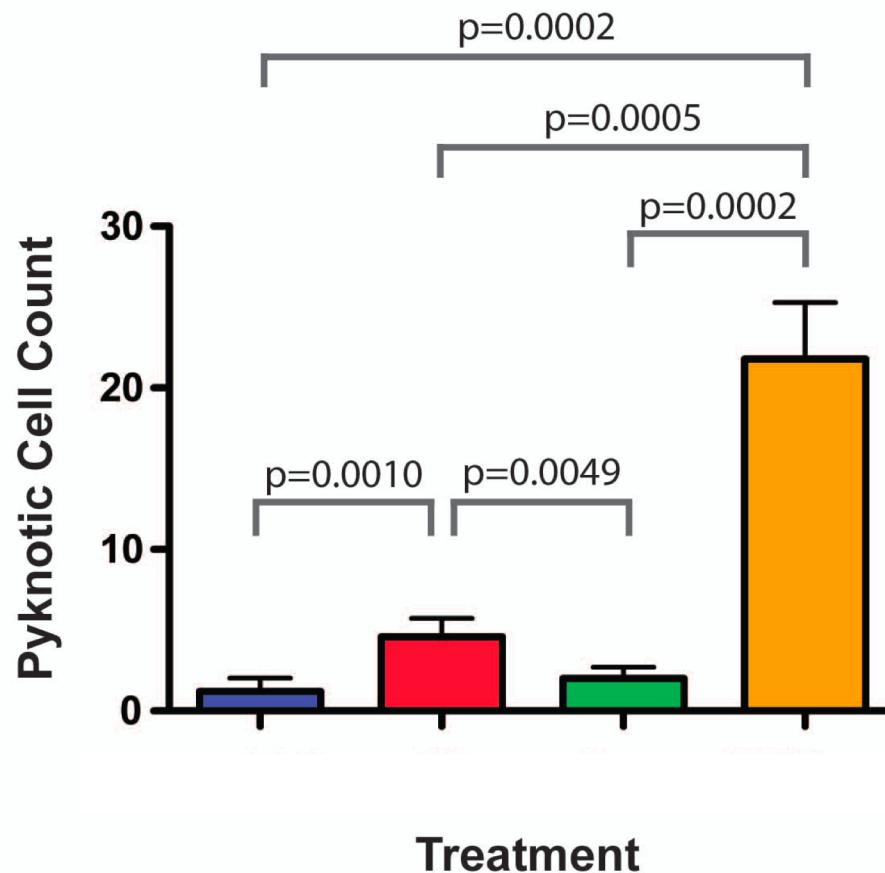
**Reviewers, Editors of “top-tier” journals,  
Grant Review Committees,  
Promotion Committees,  
and the scientific community  
repeatedly tolerate poor quality science**

**These studies typically fail at multiple levels.  
“Famous scientists” and “Top institutions” are  
repeatedly given a free-pass**

**Some examples**  
but you can find more every time you open a ‘top-tier’ journal

# Authors' Interpretation: Difference in number of dead cells

five separate fields of H&E stained paraffin tumor sections were quantified for the presence of pyknotic nuclei per square millimeter by visual inspection.

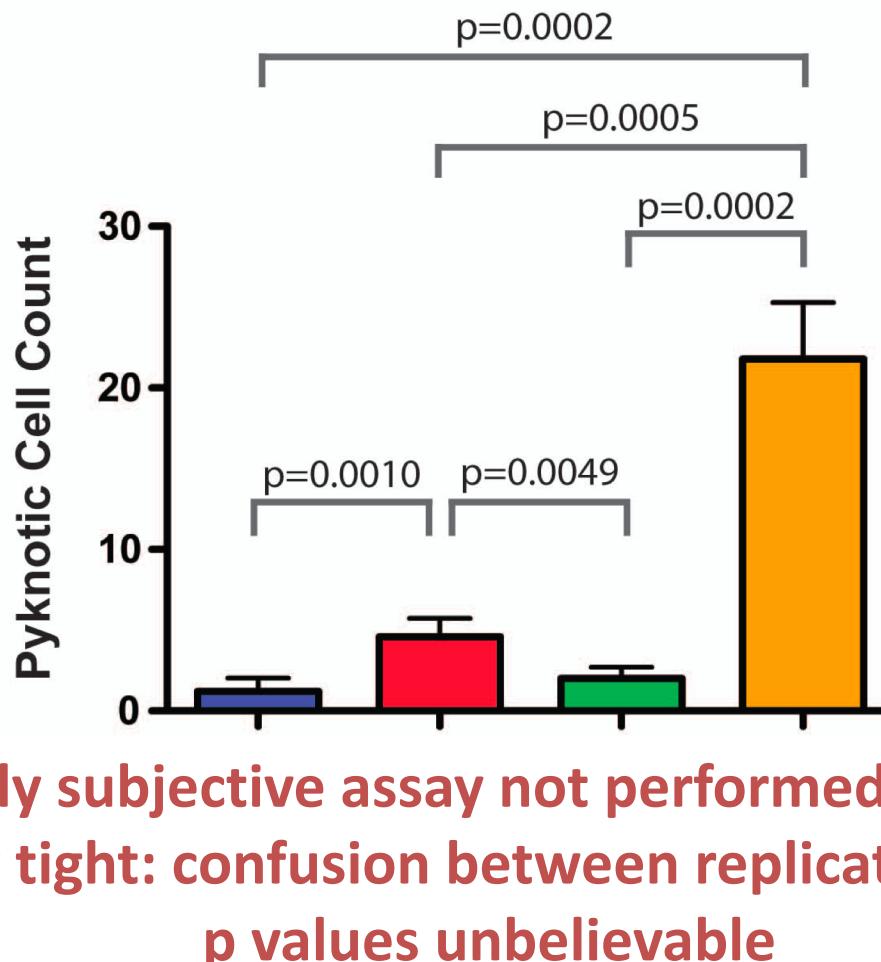


## Example #1.

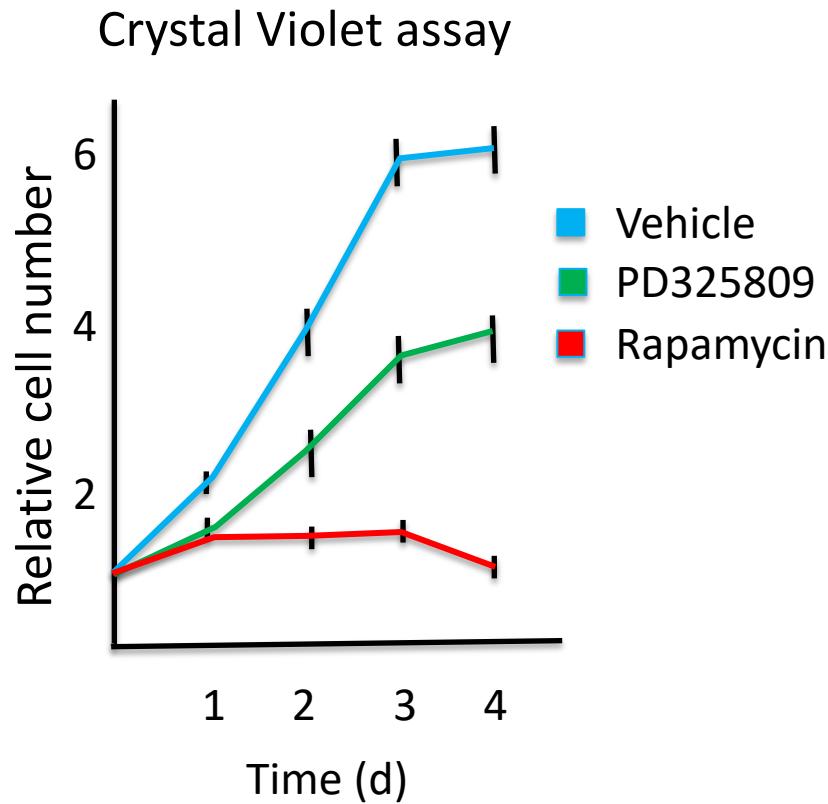
Approx 450 citations

# Authors' Interpretation: Difference in number of dead cells

Five separate fields of H&E stained paraffin tumor sections were quantified for the presence of pyknotic nuclei per square millimeter by visual inspection.

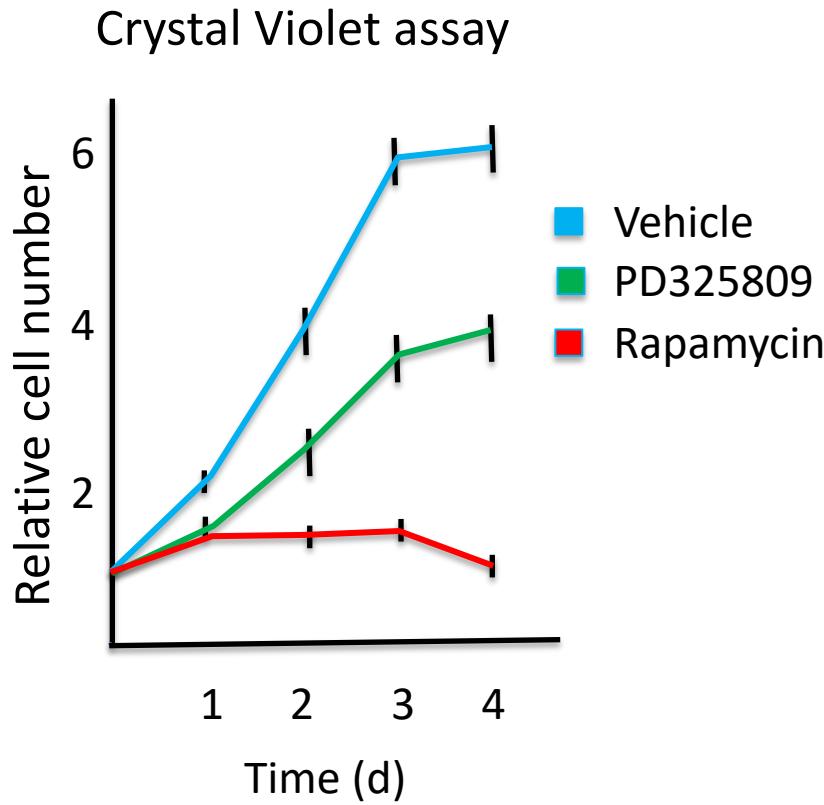


# Authors' Interpretation: Difference in cell number between 3 cell populations



Example #1.

# Authors' Interpretation: Difference in cell number between 3 cell populations



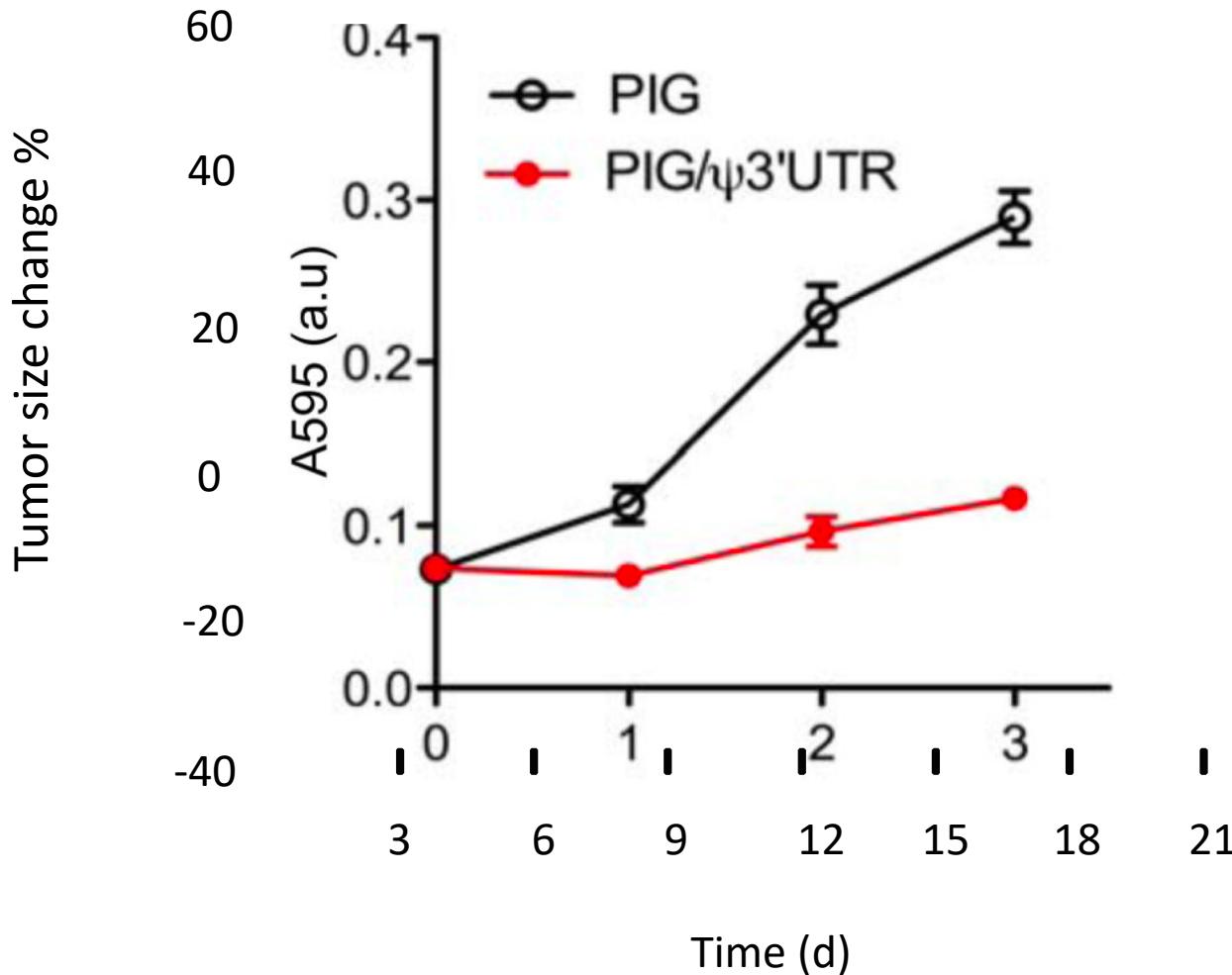
n and errors not stated

Replicates? Repeats? Not stated

Cell proliferation is an exponential, not a linear function

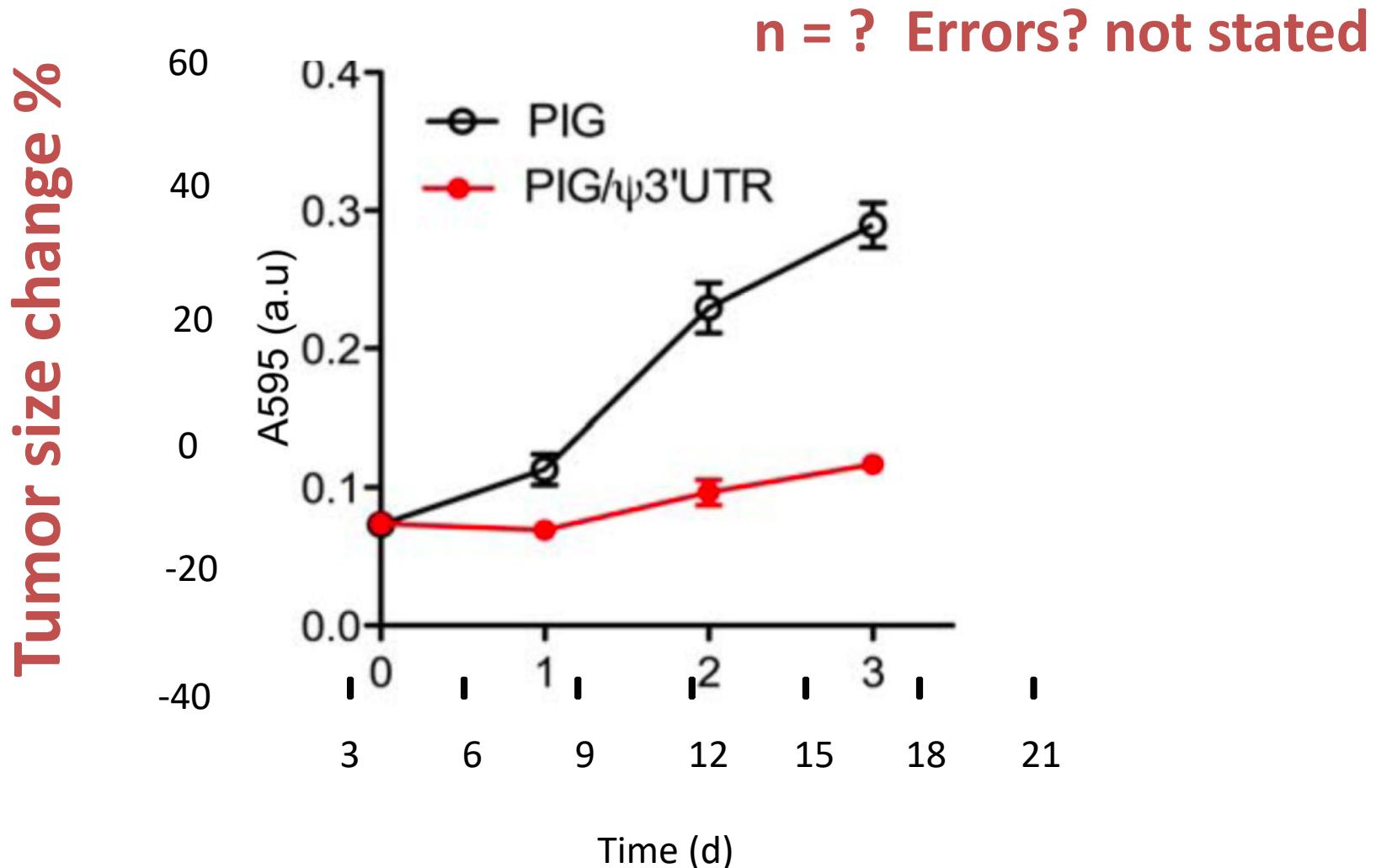
Beware results presented as “relative cell number” (what does it even mean?)

# Authors' Interpretation: Difference in tumor growth between different cohorts



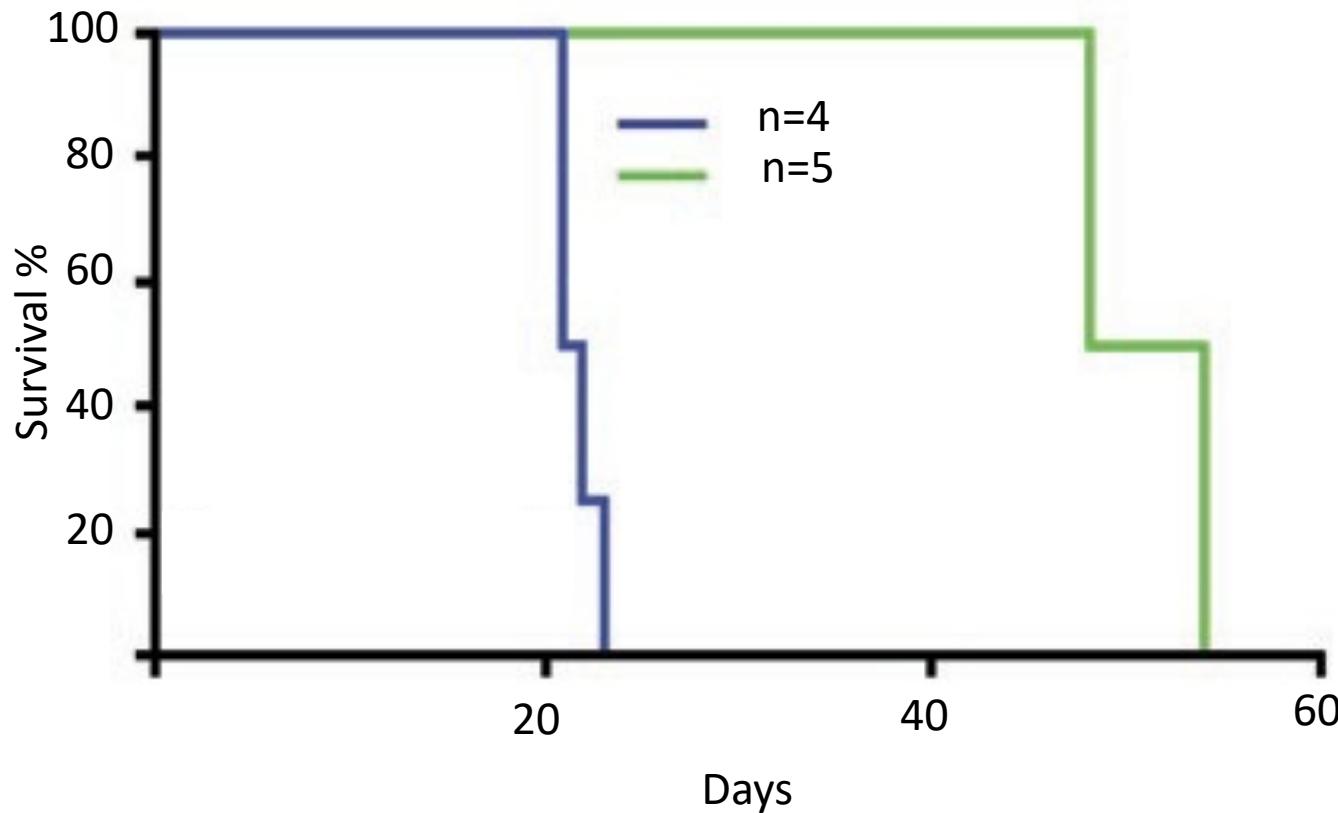
Approx 450 citations

# Authors' Interpretation: Difference in tumor growth between different cohorts



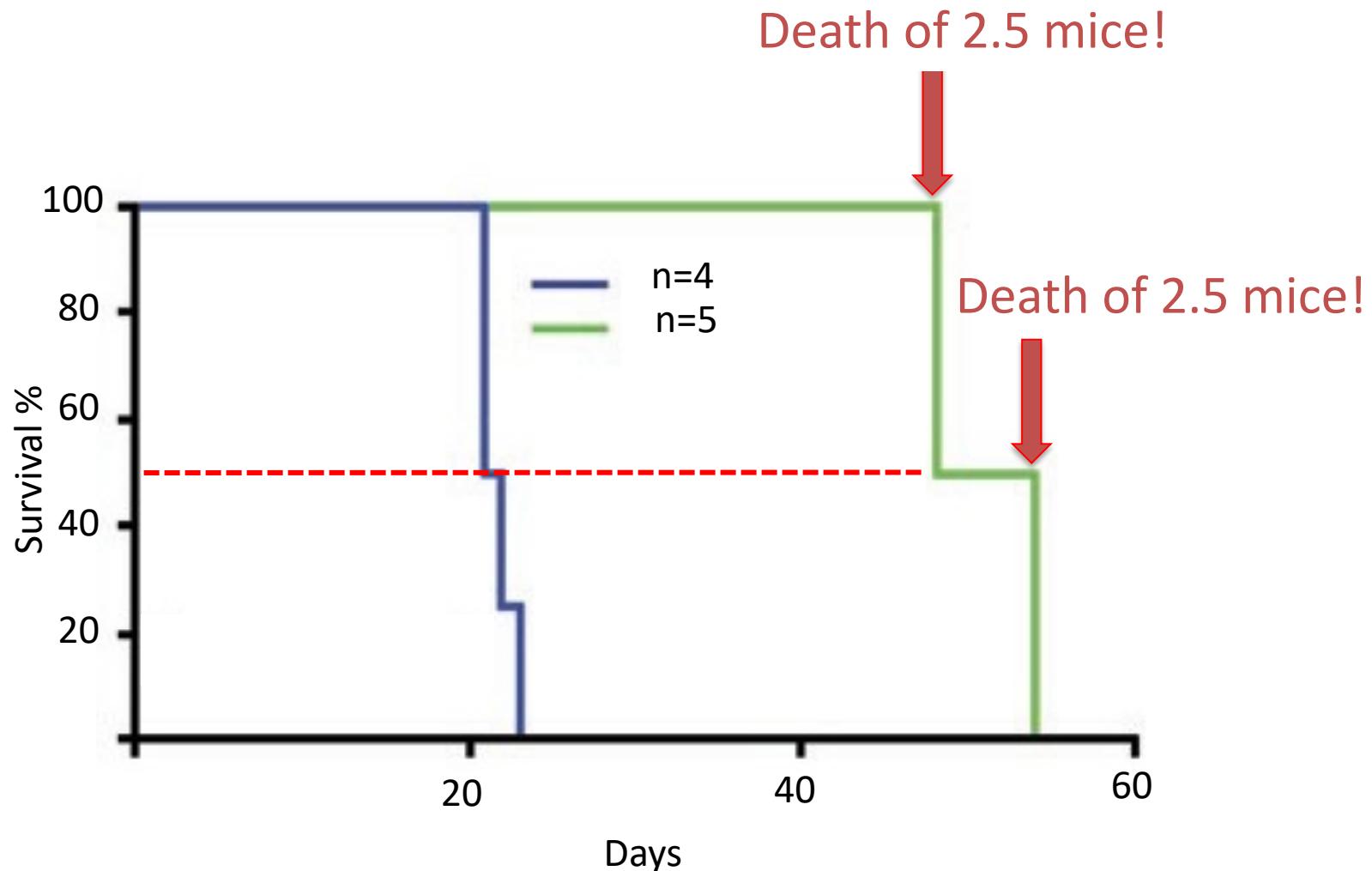
Approx 450 citations

# Authors' Interpretation: Improved mouse survival



Approx 450 citations

# Authors' Interpretation: Improved mouse survival

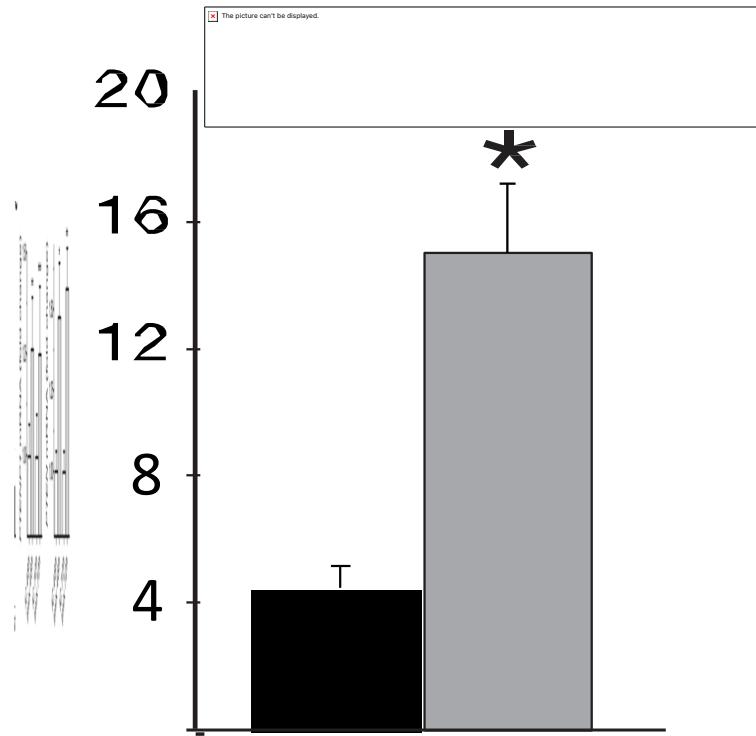
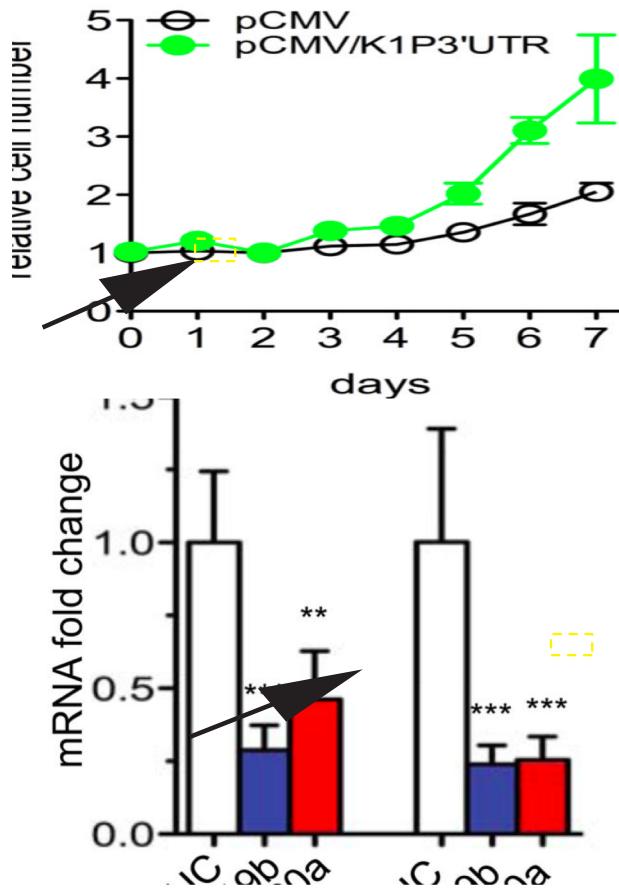


Was the paper actually read by the co-authors (n=10)?  
Reviewers? Editors? Scientific community?

Approx 450 citations

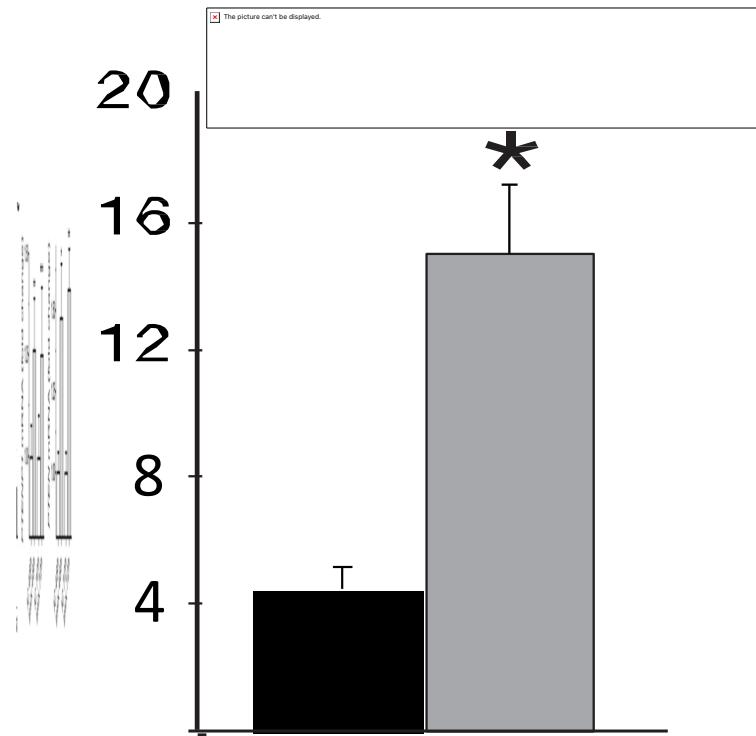
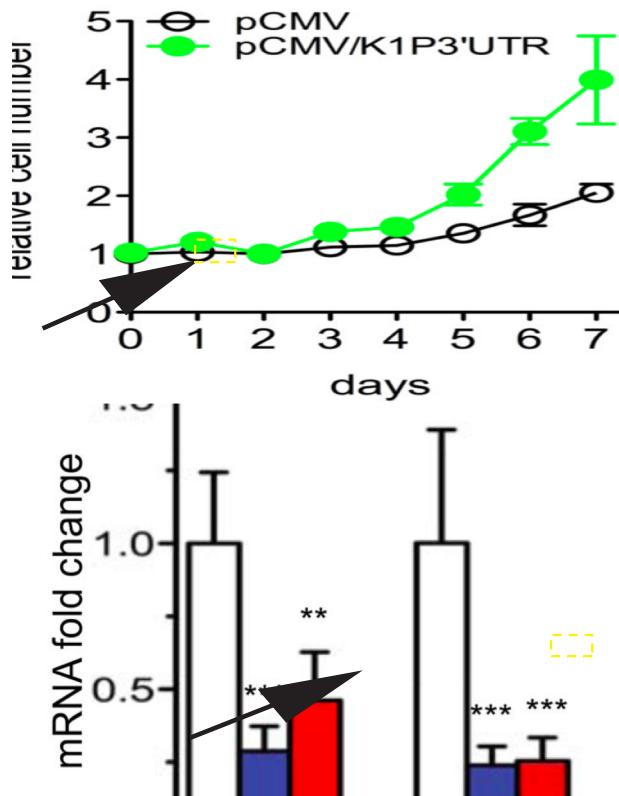
**BE SKEPTICAL**

# Authors' Interpretation: "However, metastasis was greatly enhanced..."



Example #2.

# Authors' Interpretation: "However, metastasis was greatly enhanced..."



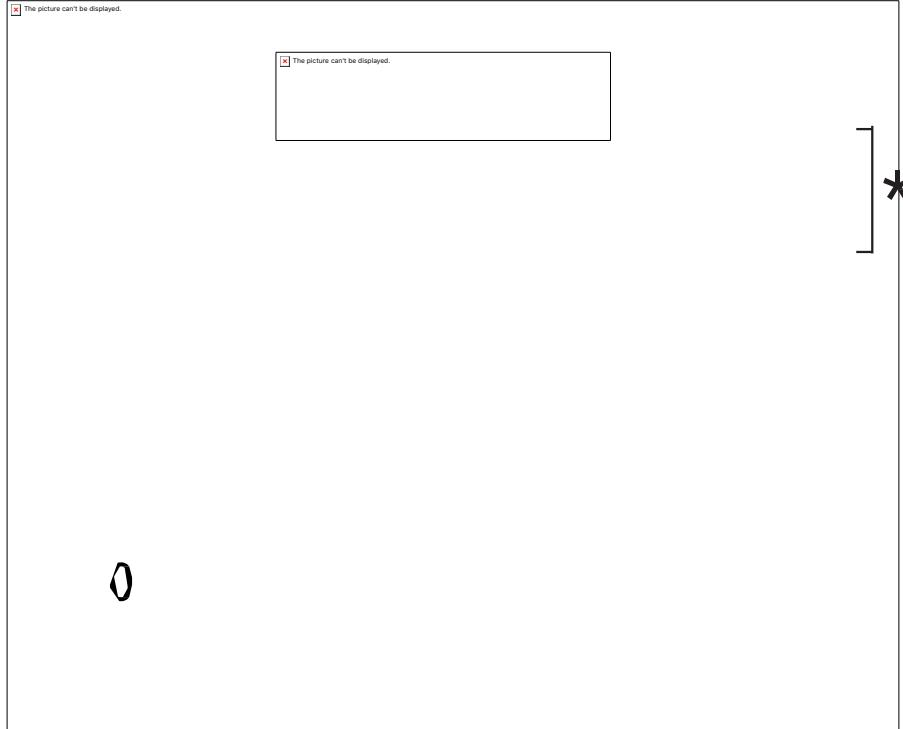
Highly subjective analysis not performed blinded

n and errors not stated

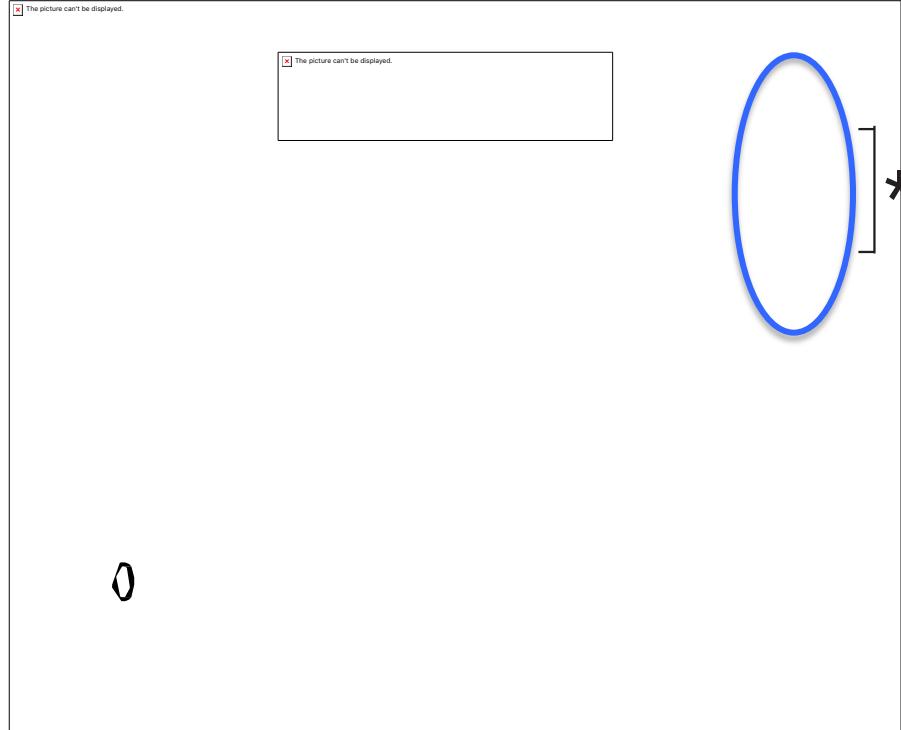
Errors are unbelievable

Beware results presented as percentage

# Authors' Interpretation: Difference in tumor growth



# Authors' Interpretation: Difference in tumor growth

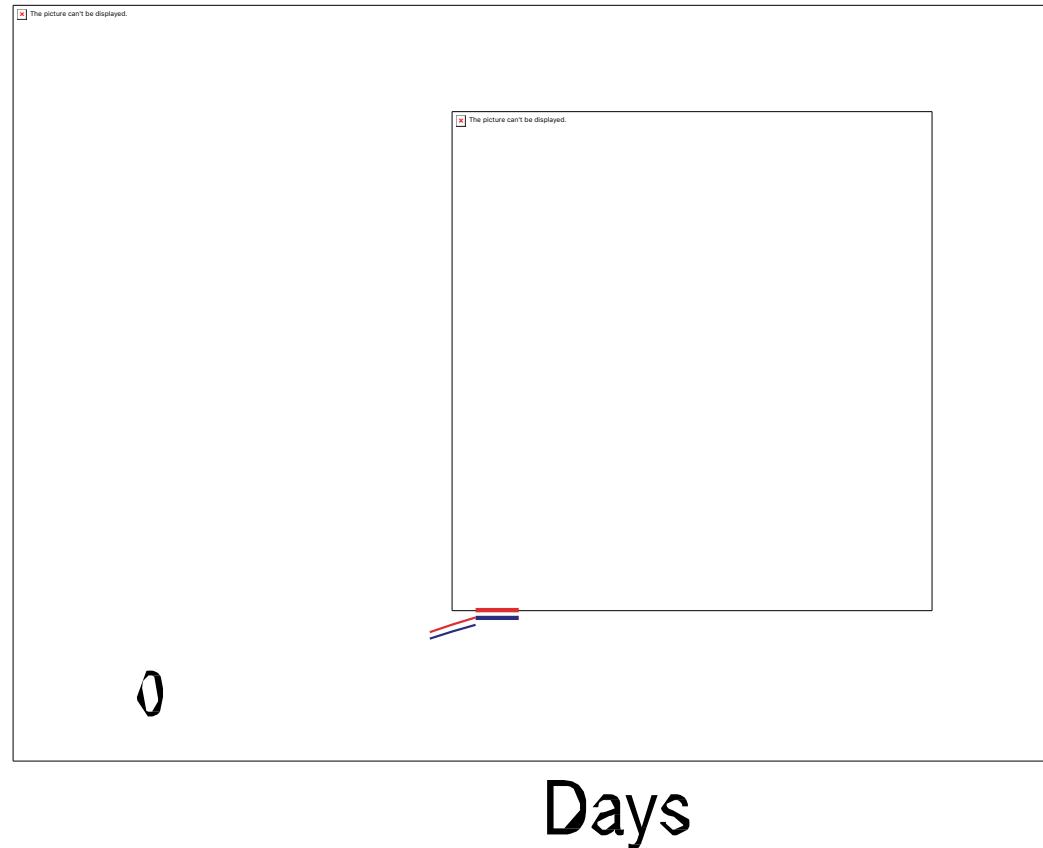


n and errors not stated.

Errors unbelievable

Beware “statistical significance” that is of no biological significance

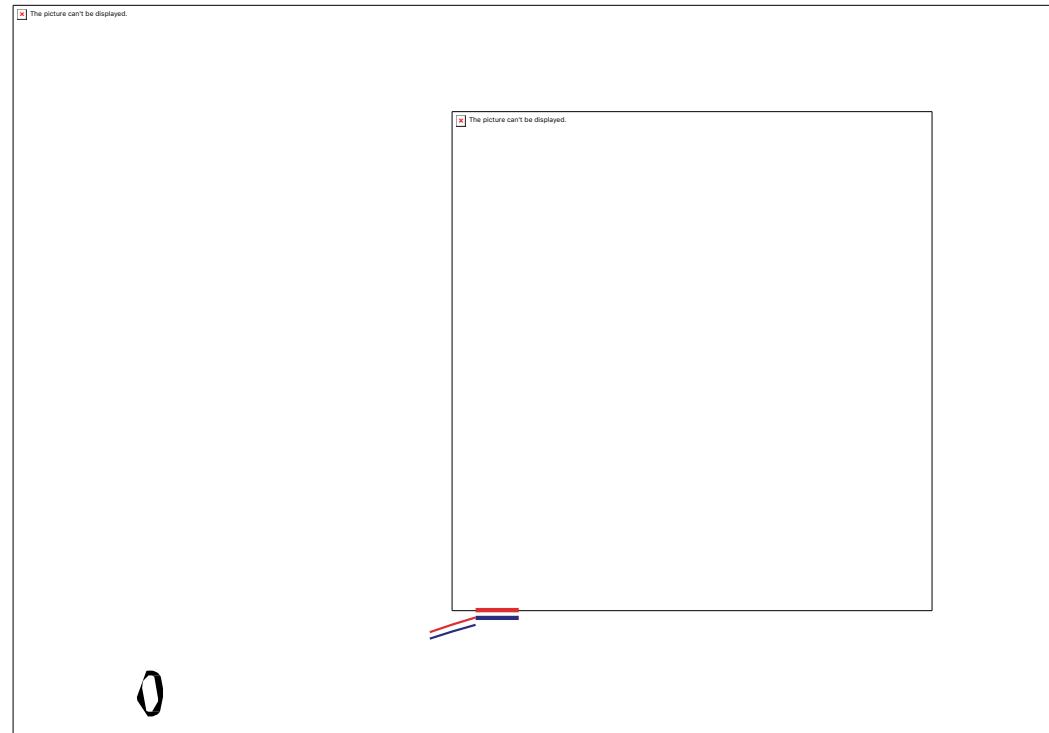
# Authors' Interpretation: Decreased tumor growth



Error bars display SEM; \* $p < 0.05$ ;

Example #2.

# Authors' Interpretation: Decreased tumor growth



0

Days

Error bars display SEM; \* $p < 0.05$

n= not stated. Errors unbelievable.

Beware “statistical significance” that is of no biological significance

Beware \* against only selected points

“...targeting of Met using PF2341066 inhibited the EMT program shift and suppressed metastasis...Although PF2341066 is also an inhibitor of ALK, quantitative RT-PCR (data not shown) and immunostaining revealed that ALK was not expressed in 4T1 tumors with or without PF2341066 treatment..”

## Example #2.

# Were positive and negative controls included/shown?

“...targeting of Met using PF2341066 inhibited the EMT program shift and suppressed metastasis...Although PF2341066 is also an inhibitor of ALK, quantitative RT-PCR (**data not shown**) and immunostaining revealed that **ALK was not expressed** in 4T1 tumors with or without PF2341066 treatment..”

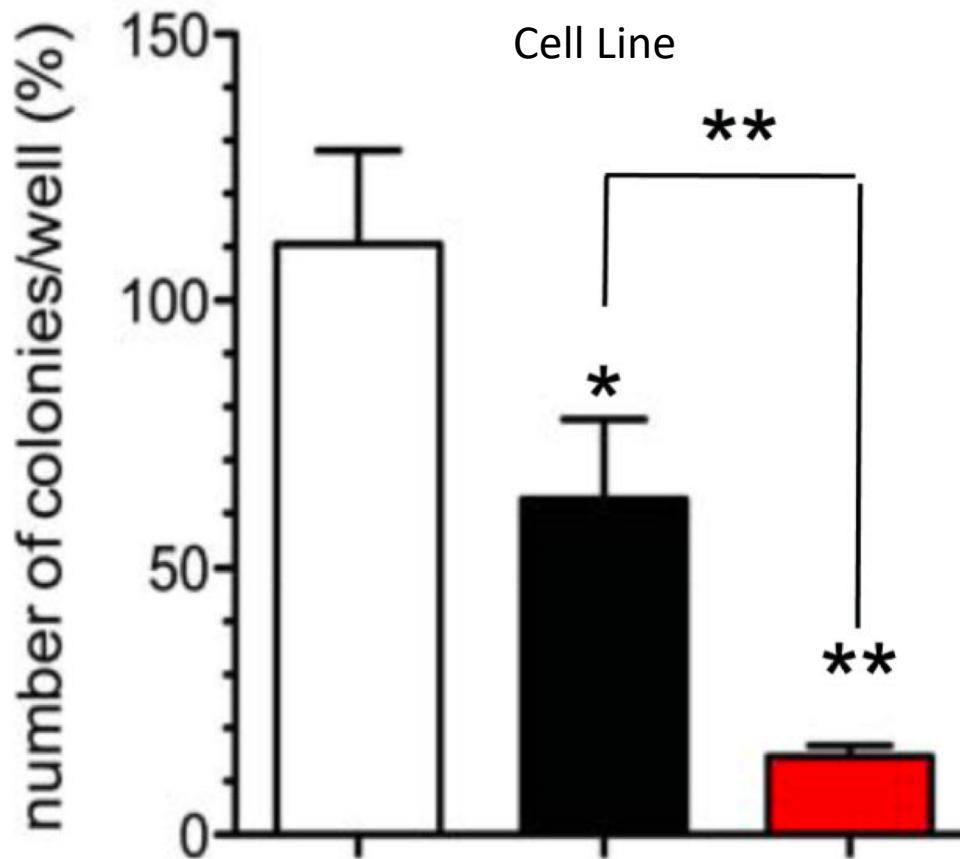
AND

Ignoring the fact that PF2341066 targets 16 kinases

**Example #2.**

**BE SKEPTICAL**

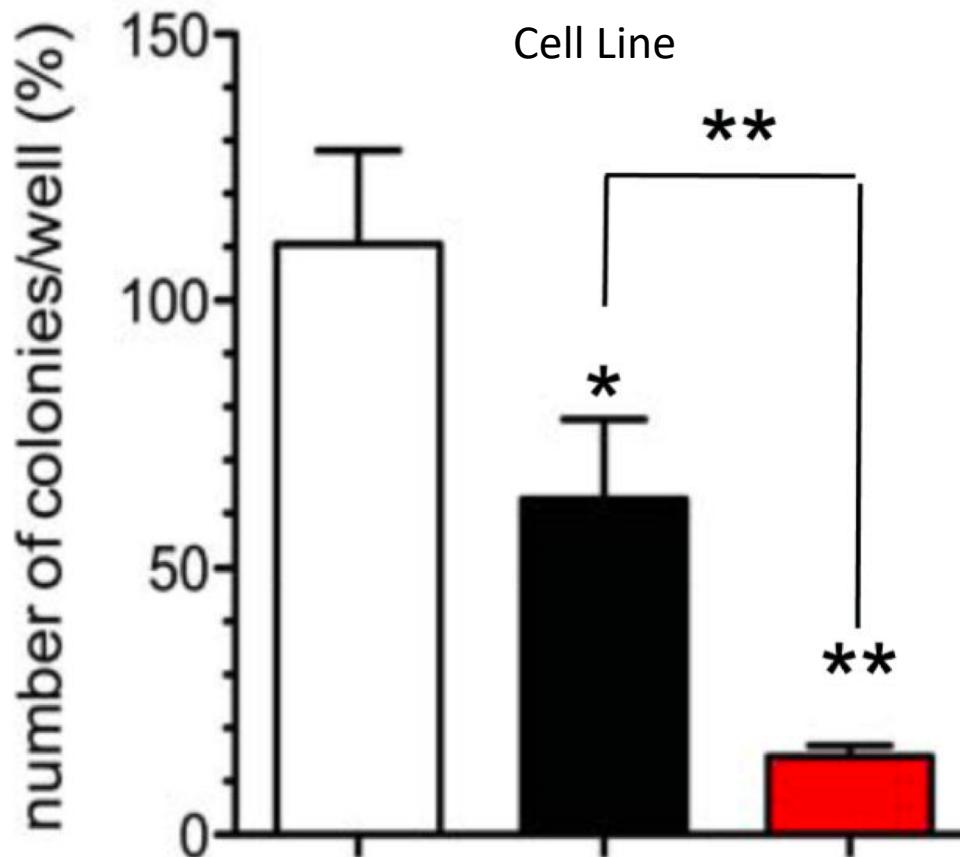
# Authors' Interpretation: Decreased colony formation



Method:  $5 \times 10^4$  cells in each well of a 6-well dish,  
colony growth scored at 3-4 weeks @ 40x magnification

Example #3.

# Authors' Interpretation: Decreased colony formation



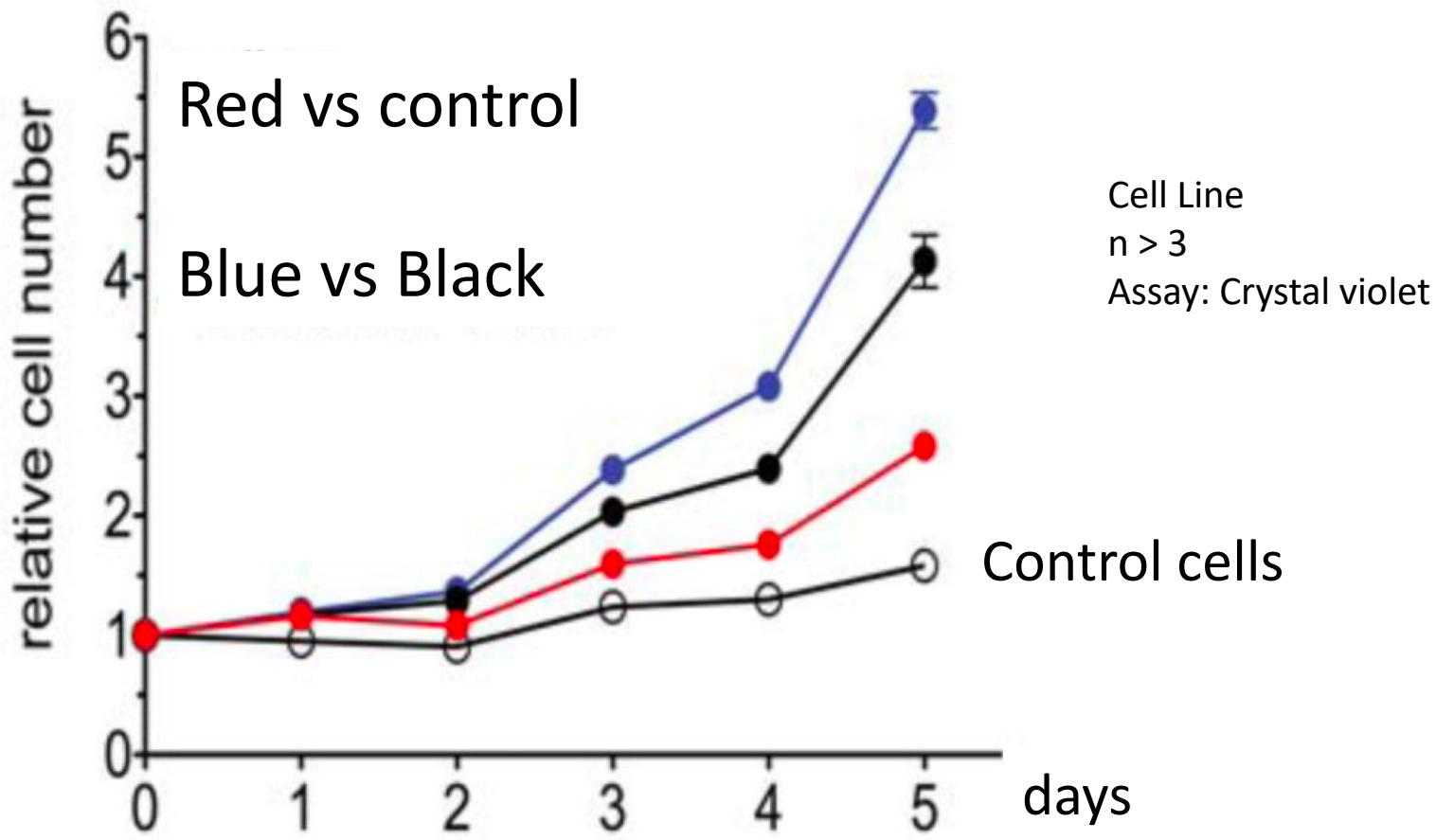
Method:  $5 \times 10^4$  cells in each well of a 6-well dish,  
colony growth scored at 3-4 weeks @ 40x magnification

Cloning efficiency  $\sim 50\%$  implies 25,000 colonies per well (35mm diam x 18mm deep)

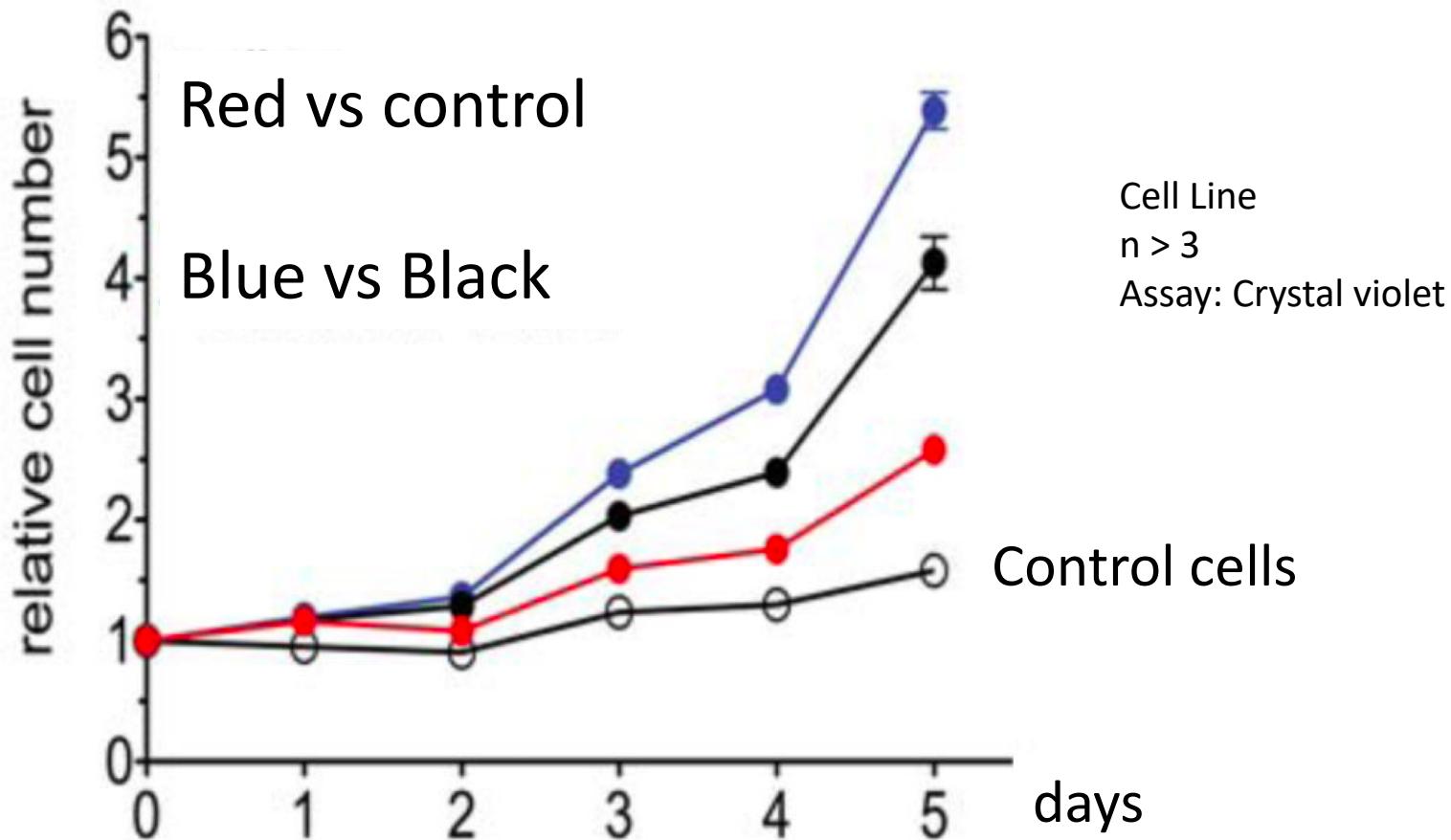
A reasonable number of 100 colonies, would require cloning efficiency of  $\sim 0.2\%$

**Beware results expressed as percentage**

Authors' Interpretation: "Transfection accelerated cell proliferation,  
suggesting that X can exert a biological activity"



Authors' Interpretation: "Transfection accelerated cell proliferation, suggesting that X can exert a biological activity"



Control cells "stationary": why is this cell line not proliferating over 5 days?

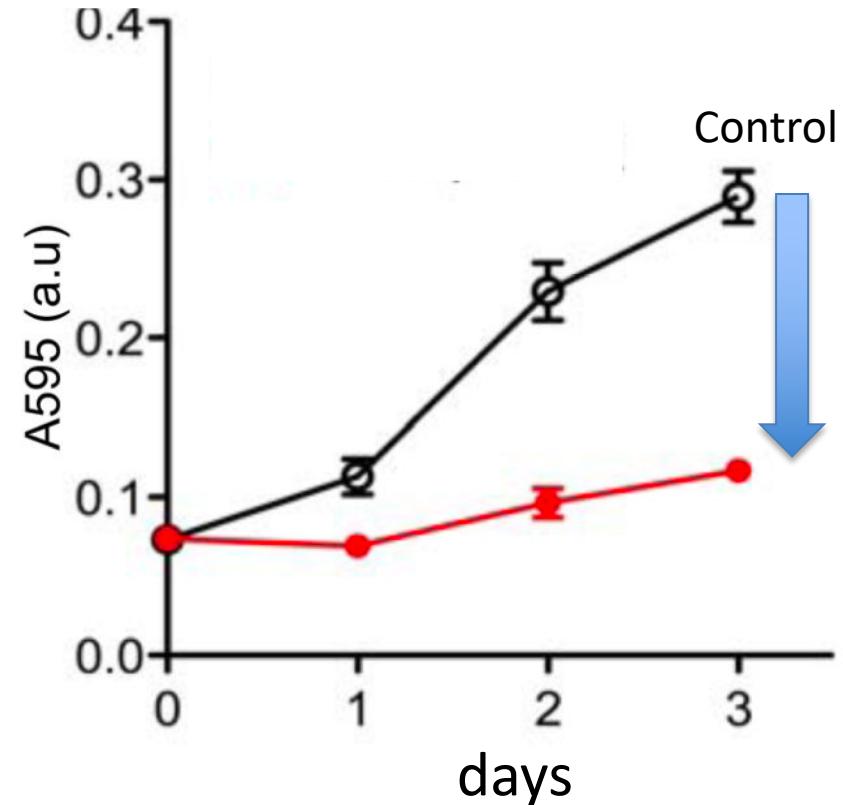
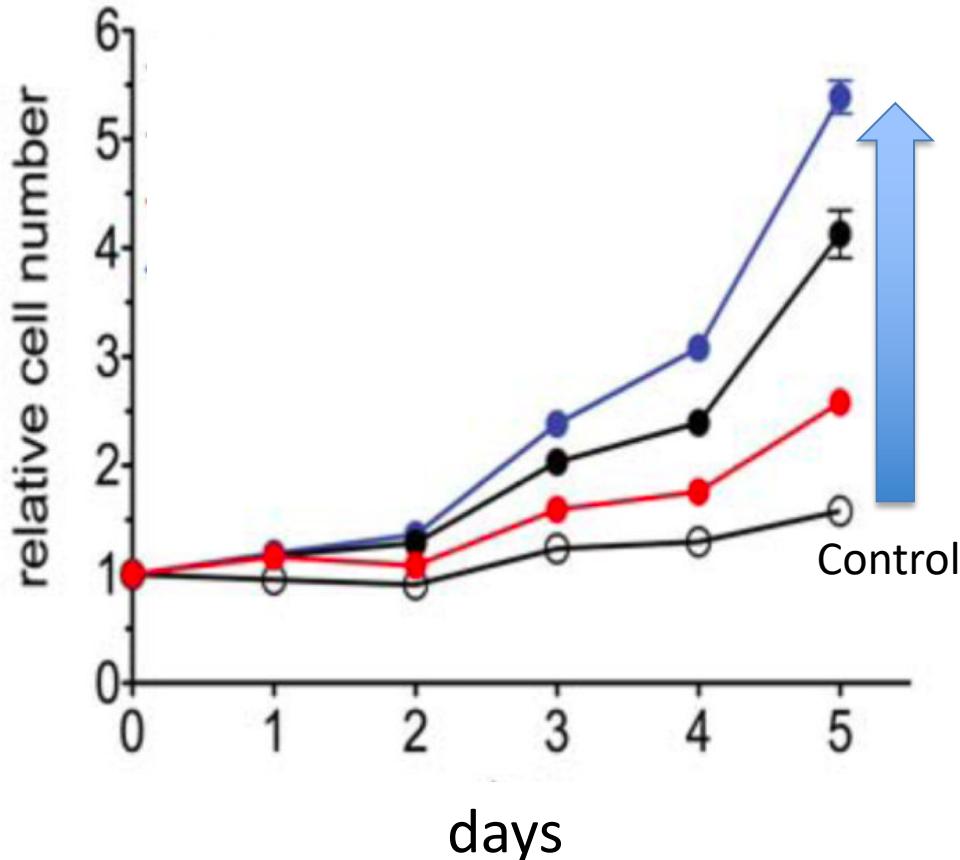
Tiny errors suggest replicate, not  $n > 3$  experiments

Linear not logarithmic Scale

Subjective assay not blinded

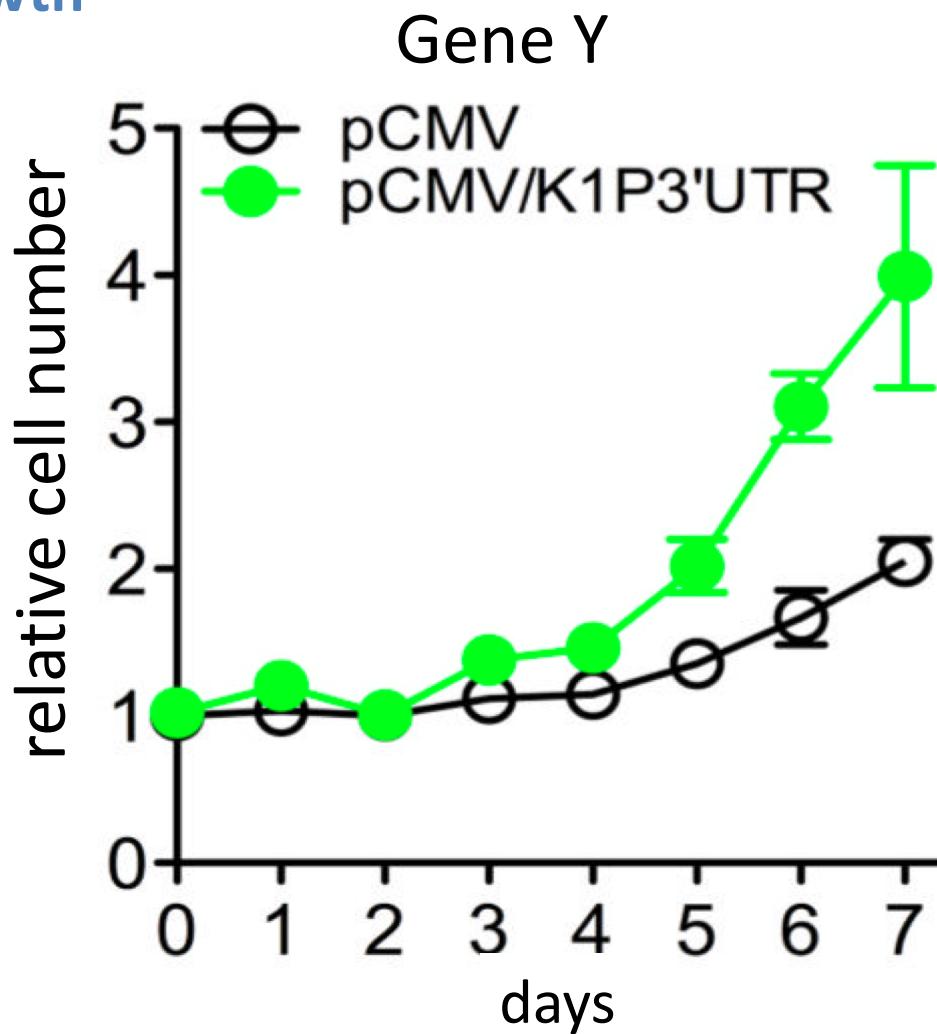
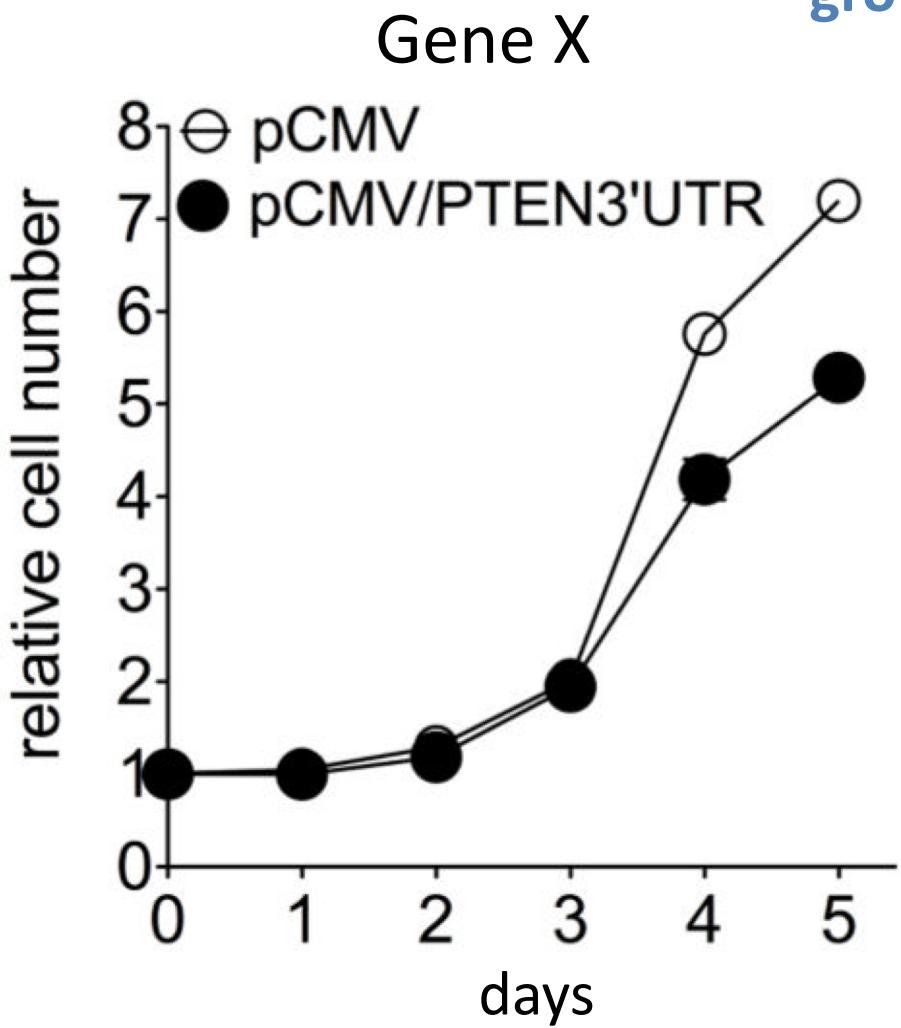
Relative, not absolute cell number

# In one Figure an increase is attributed to experimental manipulation, in another it is the converse

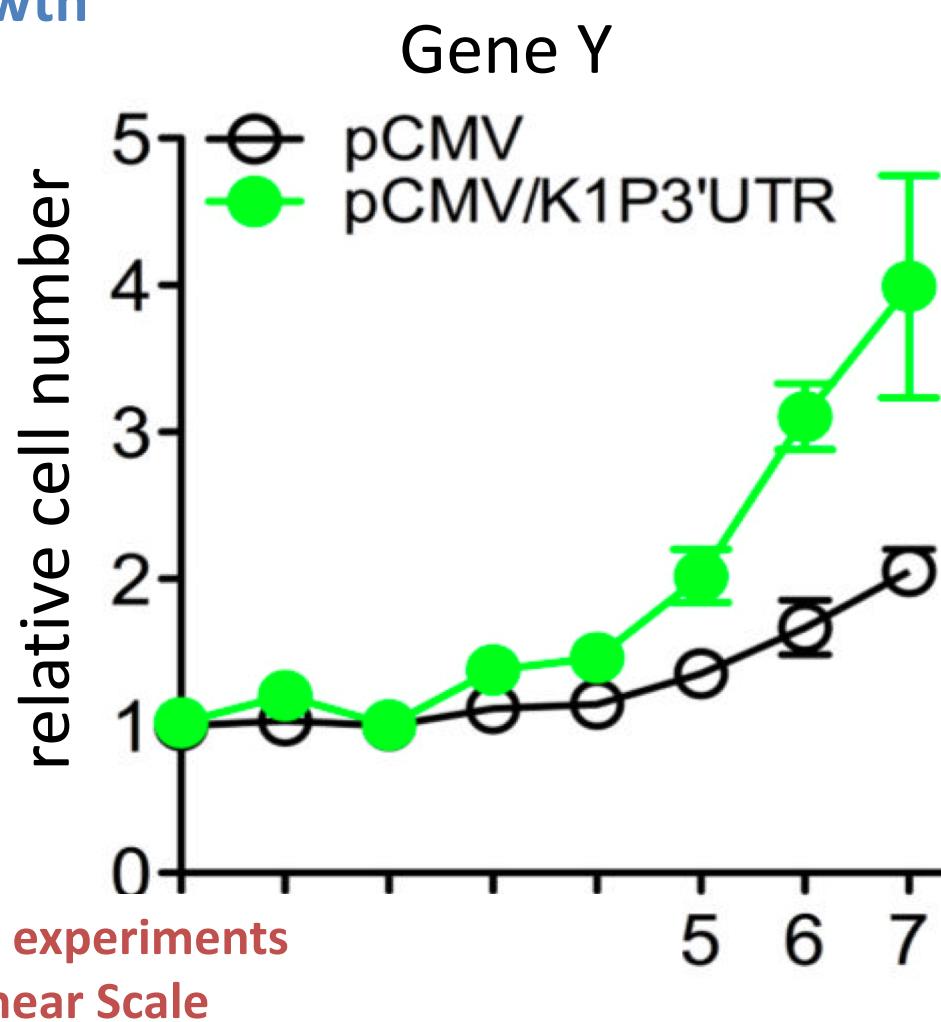
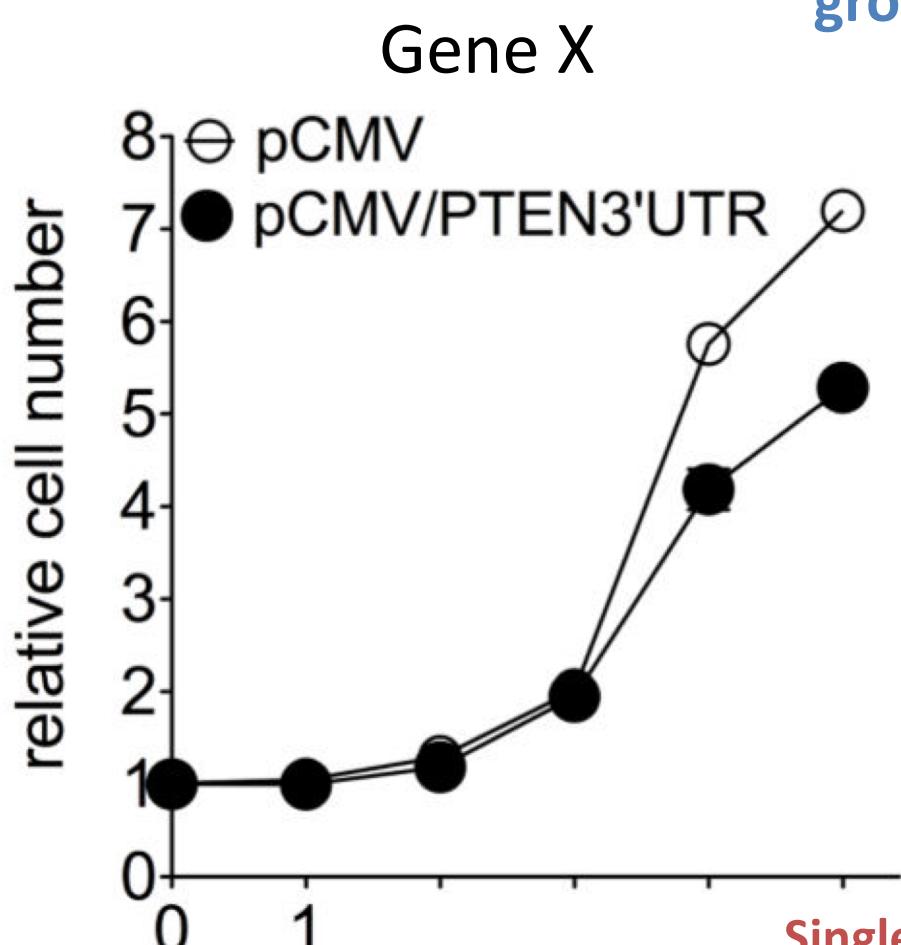


The same, control cells “stationary” for 5 days in one experiment,  
but proliferating over 3 days in the next  
....multiple examples in this paper

Authors' Interpretation: "Importantly, X overexpression was accompanied by growth inhibition...Y overexpression accelerated growth"



Authors' Interpretation: "Importantly, X overexpression was accompanied by growth inhibition...Y overexpression accelerated growth"



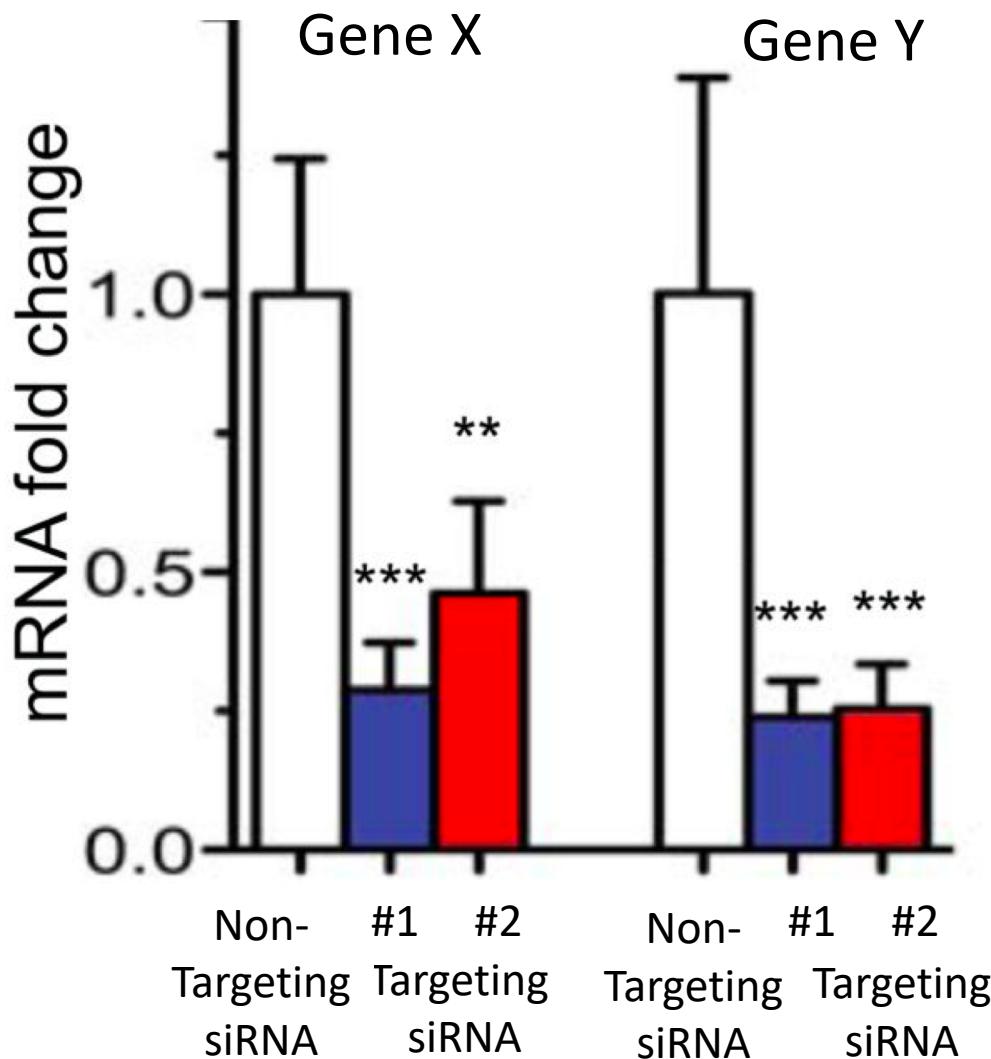
No transfection controls

Control cells growing in left panel, but not on right

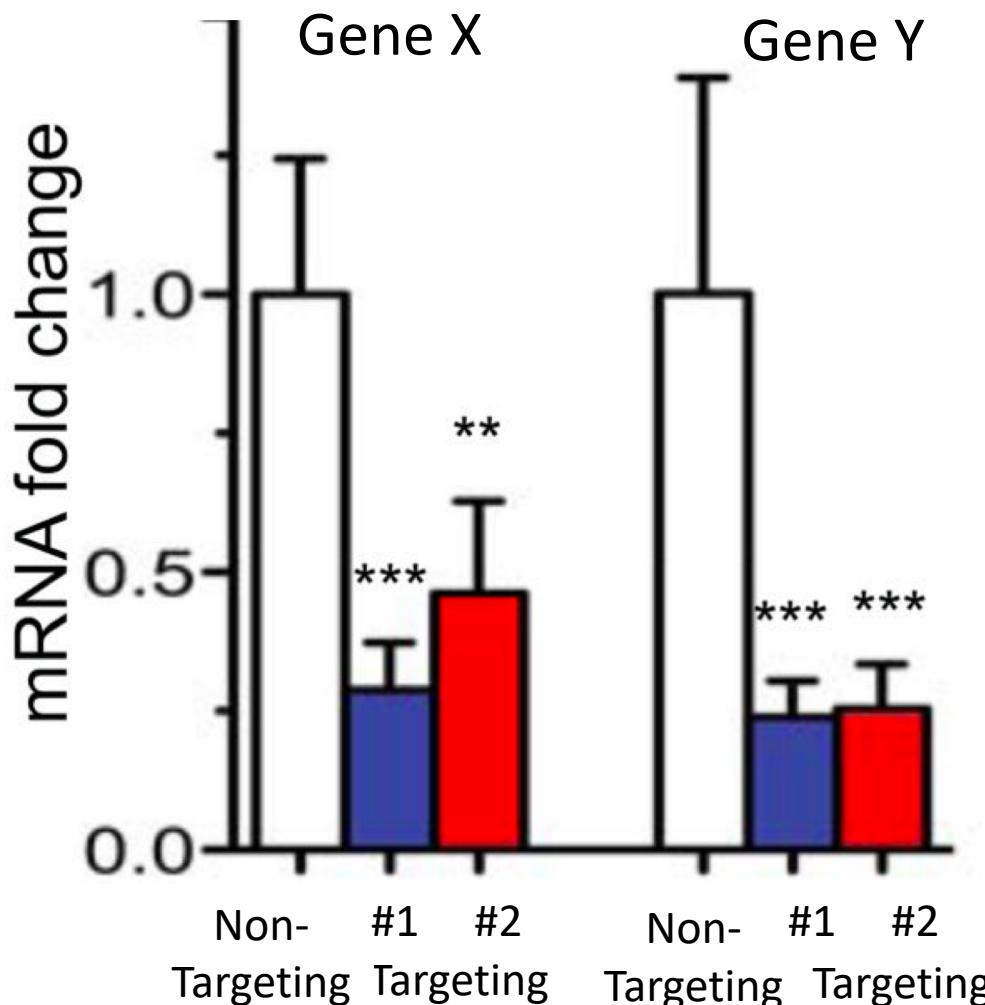
Relative not absolute cell number

Approx 800 citations

**Authors' Interpretation:** “Targeting with siRNA #1 and siRNA #2 decreased Gene X and Gene Y mRNA abundance”.



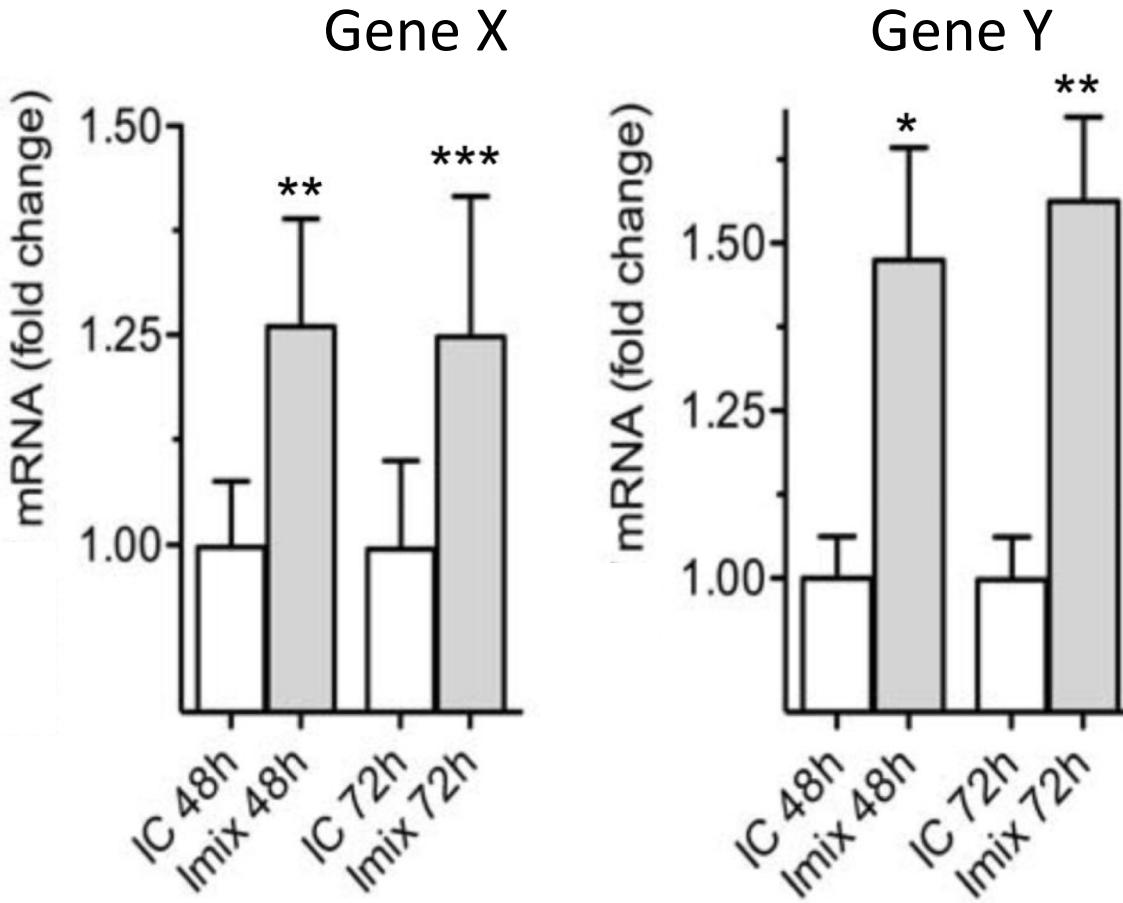
**Authors' Interpretation:** “Targeting with siRNA #1 and siRNA #2 decreased Gene X and Gene Y mRNA abundance”.



Lacks key negative control: specificity of siRNA #1 and siRNA#2  
Biological versus statistical significance of 50% decrease in mRNA

No raw data shown

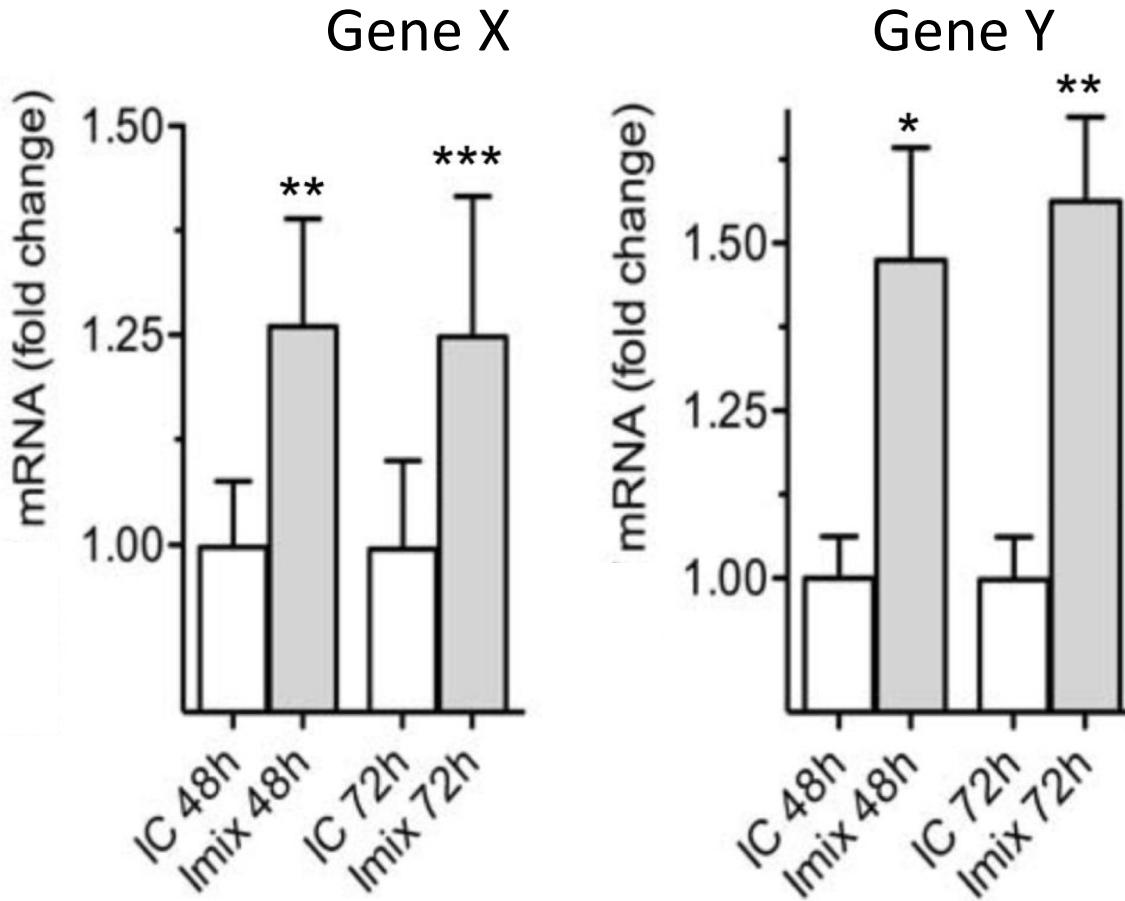
# Authors' Interpretation: A pool of micro-RNA inhibitors targeting Gene X de-repressed both Gene X and Gene Y



IC = microRNA inhibitor negative control

Imix = a pool of inhibitors of Gene X

# Authors' Interpretation: A pool of micro-RNA inhibitors targeting Gene X de-repressed both Gene X and Gene Y



IC = microRNA inhibitor negative control

Imix = a pool of inhibitors of Gene X

Lacks key negative control: specificity of inhibitors

Biological versus statistical significance? is 25% increase in mRNA relevant?

No raw data shown

## Additional concerns

- » A single cell line was studied
- » The majority of experiments were performed only once
- » No experiment stated to be performed blinded
- » In transient-transfection experiments, no data re percent of cells expressing construct
- » No evidence that transfected constructs were actually expressed
- » No evidence that antibodies were specific

**Was the paper actually read by the  
co-authors (n=6)?  
Reviewers?  
Editors?  
Scientific community?**

# There is an overwhelming bias towards publishing positive studies

Smaller, underpowered studies are more likely to yield positive results than appropriately powered studies:

Of 525 preclinical stroke studies only 2% reported a lack of effect

**BE SKEPTICAL**

# **Investigators and Institutions are primarily responsible for research integrity**

**Institutions must take greater responsibility for investigators unable to “self-censor”**

Institutions:

- take substantial overheads
- bathe in the reflected glory of their PIs
  - trumpet findings to donors and media
  - allow un-substantiated claims regarding research implications
- take insufficient responsibility for the students, post-docs
- take little responsibility for ensuring the standard of PIs
- give little recognition to excellent teachers/mentors
- do not require compliance with standard research methods
- do not monitor compliance with guidelines
- do not require/review compliance with data access
- do not review lab records

**Should demand “Good Institutional Practice”**

# Conclusions

We have a systemic problem

Our system tolerates (encourages?) these behaviors

The principal responsibility rests with the Investigator  
and their host Institution

This requires a multi-pronged approach:  
Institutions, Funders, Journals, Advocates, Press

Patients expect, and certainly deserve, more

# Recommendation:

Investigators, Institutions,

Reviewers, Funding Agencies, Consumers, Advocates, Press  
should demand:

- 1) STUDIES ARE BLINDED
- 2) All results are shown
- 3) Experiments are repeated
- 4) Positive and negative controls  
are shown
- 5) Reagents are validated
- 6) Appropriate tests are applied

