Truth Telling and Suppressed Drug Research Data

Erick H. Turner, M.D.
Senior Scholar, OHSU Center for Ethics in Health Care
Assistant Professor, OHSU Dept. of Psychiatry
Assistant Professor, OHSU Dept. of Pharmacology
Staff Psychiatrist, Portland VA Medical Center
Former FDA Medical Officer

Science and the pursuit of truth

• From Latin: “scientia” = knowledge
• A scientific theory must be testable. It must be possible in principle to prove it wrong.
• Experiments are the sole judge of scientific truth.
• Scientific method:
  – Observations
  – hypothesis/theory
  – experiment (test)
  – revision of theory.

http://www.astronomynotes.com/scimethd/s2.htm
Cycle of knowledge in medicine (EBM)

Hypothesis (protocol) → Test hypothesis (clinical trial) → Observations (trial results) → Communicate observations (journal publication) → Hypothesis (protocol)

Communication of Results

• Critical to this process is making every relevant aspect of research publicly available, which allows ongoing review and repeating of experiments and observations by multiple researchers operating independently of one another.

Publication bias

- **Definition:** Occurs when...
  - the publication of research results depends on their nature and direction
  - authors are more likely to submit, or editors accept, positive than null (negative or inconclusive) results

- **Types**
  - Delayed publication (negative results)
  - Multiple publication (positive results)
  - Selective publication
    - Aka "the file drawer problem"
    - Can apply to:
      - Trials per se
      - Outcomes within those trials

- **Domains (drugs)**
  - Safety
  - Efficacy

My disillusionment

- **Before FDA**
  - Journal articles reported only positive findings

- **At FDA**
  - Lots of negative trials
  - "What gives, boss?"
  - Learned FDA review documents publicly available
    - Older drugs -- "paper FOIA"
    - Newer drugs -- on FDA website
Overview of our study


- Cohort of studies for one drug class
  - 12 antidepressants starting with Prozac
  - All studies registered with FDA
- FDA reviews → identify all studies
- Track each study into published literature
- Two questions:
  - Was the study published?
  - If published, how did the published results compare with the FDA results?

**FDA-positive studies (N=38)**

<table>
<thead>
<tr>
<th>Drug</th>
<th>Publication Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>nefazodone</td>
<td>Published, agree with FDA</td>
</tr>
<tr>
<td>citalopram</td>
<td>Published, agree with FDA</td>
</tr>
<tr>
<td>escitalopram</td>
<td>Published as positive, in conflict with FDA</td>
</tr>
<tr>
<td>bupropion SR</td>
<td>Published, agree with FDA</td>
</tr>
<tr>
<td>paroxetine</td>
<td>Published, agree with FDA</td>
</tr>
<tr>
<td>paroxetine CR</td>
<td>Published as positive, in conflict with FDA</td>
</tr>
<tr>
<td>fluoxetine</td>
<td>Published as positive, in conflict with FDA</td>
</tr>
<tr>
<td>duloxetine</td>
<td>Published, agree with FDA</td>
</tr>
<tr>
<td>mirtazapine</td>
<td>Published as positive, in conflict with FDA</td>
</tr>
<tr>
<td>sertraline</td>
<td>Published as positive, in conflict with FDA</td>
</tr>
<tr>
<td>venlafaxine</td>
<td>Not published</td>
</tr>
<tr>
<td>venlafaxine XR</td>
<td>Not published</td>
</tr>
</tbody>
</table>
Failed studies (N=12)

- Published, agree with FDA
- Published as positive, in conflict with FDA
- Not published

<table>
<thead>
<tr>
<th></th>
<th>CH104-306</th>
<th>BCT-HD-02</th>
<th>03-002</th>
<th>448</th>
<th>MMQA-A</th>
<th>MMAT-A</th>
<th>103</th>
<th>KDA-303</th>
<th>307</th>
</tr>
</thead>
</table>

- nefazodone
- citalopram
- escitalopram SR
- bupropion SR
- paroxetine
- paroxetine CR
- fluoxetine
- duloxetine
- mirtazapine
- sertraline
- venlafaxine
- venlafaxine XR

Negative studies (N=24)

- Published, agree with FDA
- Published as positive, in conflict with FDA
- Not published

<table>
<thead>
<tr>
<th></th>
<th>CH104-007</th>
<th>693-007</th>
<th>98006</th>
<th>80001</th>
<th>62-001</th>
<th>62-007</th>
<th>68</th>
<th>80-001</th>
<th>315</th>
</tr>
</thead>
</table>

- nefazodone
- citalopram
- escitalopram SR
- bupropion SR
- paroxetine
- paroxetine CR
- fluoxetine
- duloxetine
- mirtazapine
- sertraline
- venlafaxine
- venlafaxine XR
Overall - Studies

- Published, agree with FDA
- Published as positive, in conflict with FDA
- Not published

Positive
N=37 (97%)

Failed
N=38 (51%)

Not positive
N=36 (49%)

N=1 (3%)

Overall - patients in studies

- Published, agree with FDA
- Published as positive, in conflict with FDA
- Not published

Positive
N=7,075 (99%)

Failed
N=5,409 (43%)

Not positive
N=3,369 (62%)

N=197 (4%)

N=80 (1%)
Effect size (ES) approach

• P values:
  – Tell you whether the effect is zero or not
  – No information on how big the effect is
  – Subject to manipulation by $N$
    • $N$ too small -- important finding can be “NS”
    • $N$ too big -- trivial finding can be significant
  – Statistical significance $\neq$ clinical significance

• Effect size (ES)
  – Tells 
    magnitude of effect (clinical significance)
  – Not influenced by $N$
  – Meta-analysis: combine results of different studies to get the “big picture”

• $ES_{published}$ vs. $ES_{FDA}$

$ES_{published}$ vs. $ES_{FDA}$: column graph
Eleven pigs with lipstick

Game Time

Pick the piggy!

- 3 pairs of peer-reviewed journal articles
  - Same author for each pair

- Within each pair, study circled main result:
  - 1 agrees with the FDA
  - 1 is a pig with lipstick

- Write down your decision on the paper
- We’ll tally votes then check FDA result

Goldstein 2002
J Clin Psychiatry

- $P_{\text{published}} = 0.009$ (highly signif)
- $P_{\text{FDA}} = 0.14$ (NS)
- Word “primary” used (unusual)
  - Yes re rating scale
  - Not w/ statistical method
- Primary method for handling data from dropouts
  - Lilly proposed MMRM
  - FDA said no -- LOCF primary

Mendels 1993
Psychopharm Bull.

- $P_{\text{published}} < 0.05$ (dose of 150-200mg/day)
- $P_{\text{FDA}} = 0.14$ (NS)
- Emphasized signif OC results
- Primary analysis (LOCF)
  - Means given
  - no P value
  - No statement re lack of signif
- LOCF scores were used to get signif “trend analysis”
  - Didn’t mention lack of dose-response with OC
    - Fig: 25-mg > 75mg
Rickels 1986
Current Therapeutic Research

- $P_{\text{published}} < 0.10$
  - “…should be interpreted in the light of the small sample size available.”
  - Thus suggests it would have been significant if only $N$ had been larger
- Reported $P$ value is 1-tailed
  - Convention in psychiatric trials is to use 2-tailed $P$ values
  - $P_{1\text{-tailed}} < 0.10$ same as $P_{2\text{-tailed}} < 0.20$ (looks much worse)
- $P_{\text{FDA}} = 0.50$ (worse still)
- Fluoxetine actually did numerically worse than placebo
- Used observed cases method (“OC”, completers only; dropouts ignored)
  - Discussion: “Because of the slow onset of action, endpoint analyses were considered not appropriate for this data set.”

How is this possible?

Why the discrepancies between the FDA and journal articles?
Information flow at the FDA

• **Before** the study
  – Sponsor notifies FDA of intent to do study (registration)
  – FDA gets sponsor’s full protocol up front
    • Primary outcomes declared -- basis for “win” or “loss”
      – Scale
      – “Devilish” details re statistical analysis
      – Secondary outcomes = exploratory only
  
• **After** the study
  – Receives study reports (all registered studies, summary statistics + raw data)
  – Re-analysis
    • prespecified methods → replication
  – Written up in review document
    • Review posted on web if drug approved

Information flow - peer-reviewed system

• Conceive study
• Write protocol
  – Collect lots of data -- many scales, etc.
    • Danger of multiplicity (shotguns seldom miss)
  – One should be primary *(the hypothesis)*
• Get regulatory approval
• Do the study
• Analyze the data
  – Try out alternate methods p.r.n.
  – Torture data until it confesses

• Then…
Write it up!

• **If** you choose to (nonpublication)
• **When** you choose to (delayed publication)
• **How** you choose to (outcome reporting bias)
  – Journals seldom ask for the original protocol
  – “Honor system”
  – System invites HARKing

System invites HARKing
Selective reporting of outcomes within studies

**Hypothesizing**
• Decide—enlightened by the data—
  which methods you used were flawed / valid

**After the**
• Decide what results do / don’t merit emphasis in the manuscript

**Results are**
• Look like a genius!

**Known**
• Like betting on a horse race *after* the race

Recipe for success

1. Find out which horse wins
2. Place your bet
3. You win!! (You can’t lose!)

Incentives for HARKing & selective publication

- Hindsight bias (“I knew it all along!”)
- Money
  - Increase drug sales
  - Grant money
    - Pilot data must “show promise” to impress review committees to invest in proposed study
- Hot stuff (journals & press)
- Narrative flow / satisfying story line
- Ego (“I was right!”)
- System practically begs you to do it
  - Few journals ask for original protocols
What journals should do

• Protocol as covenant / contract with society
• Make protocol available to...
  – Reviewers and editors
  – Readers

Consequences of publication bias

• Efficacy
  – “Grade inflation” -- Lake Wobegone effect -- all the treatments are above average
• Safety
  – Harms concealed or played down
  – Vioxx, Avandia, many more
• Risk-benefit ratio skewed
  – Overenthusiasm about treatment options
• Scientific method undermined
  – Evidence b(i)ased medicine
  – It’s not the truth!
Pervasiveness of publication bias

• Industry-sponsored drug trials
• Government-sponsored trials
  – NIH-sponsored (Dickersin & Min, Online J Curr Clin Trials. 1993 Apr 28;Doc No 50)
  – Canadian Inst. of Health Res. (Chan, et. al. CMAJ. 2004 Sep 28;171(7):735-40)
• Other areas of medicine
  – Cancer research (Kyzas, et al 2007; Ramsey & Scoggins 2008)
• Dental research (Scholey JM, Harrison JE. Evid Based Dent. 2005;6(3):58-61)
• CAM trials (Ernst E. J Clin Epidemiol. 2007 Nov;60(11):1093-4)
• Social sciences

Policy fixes and limitations

• Registration
  – ICMJE (2005) - must register trial if you want to publish
  – But publishing is optional!
  – Can’t verify results
    • Declaring primary outcomes is optional
    • Primaries can be vague (“depression rating scale”) → HARKing
• Posting of results by pharma companies
  – Only go back so far (often 2004)
  – Awareness? Doctors still rely on peer-reviewed literature
  – Not a disinterested 3rd party
• FDAAA of 2007
  – Registry and results database (clinicaltrials.gov)
  – Only affects future drugs!
    • Today’s drugs grandfathered in & not covered
What about all those drugs on the market now?

Today’s drugs are not going away soon

Top 200 Drugs: [www.drugtopics.com](http://www.drugtopics.com)

*Turner, et al, Science Oct 2008*

---

That goldmine of clinical trial data we’re still ignoring

*Public Library of Science Medicine, Oct 2004*
An Ethical Framework for Clinical Research

• “Develop generalizable knowledge to improve health and/or increase understanding of human biology”
• Obligation to produce unbiased knowledge (truth)

Emanuel EJ et al. JAMA. 2000;283:2701-11

Questions to Consider

• Why not just leverage our FDA data?
  – “Whose tax money is it, anyway?”
    • (Paraphrasing John Santa)
• Who owns the data (knowledge)?
  – Drug company?
    • “trade secrets” and “confidential information”
      – Exemption in FOIA
    – subjects, investigators, funders, public
• Should informed consent forms act as covenant with subjects that their data will be disclosed truthfully?