

# Clinical Research Data Sharing Practices and Attitudes

Andrew J. Vickers

Department of Epidemiology and Biostatistics  
Memorial Sloan-Kettering Cancer Center



# Cancer Data? Sorry, Can't Have It

By ANDREW VICKERS

Not long ago, I asked a respected cancer researcher if he could send me raw data from a trial he had recently published. He refused. Sharing data would make the study team members "uncomfortable," he said, as I might use this to "cast doubt" on their results.

I'd heard this before: as a statistician who designs and analyzes cancer studies, I regularly ask other researchers to provide additional information or raw data. Sometimes I want to use the data to test out a new idea or method of statistical analysis. And knowing exactly what happened in past studies can help me design better research for the future. Occasionally, however, there are statistical analyses I could run that might make an immediate and important impact on the lives of cancer patients.

A few years back, a study was published showing that a certain drug could prevent one type of cancer. The problem was that the drug didn't work very well and had some side effects, so almost no one used it. At the same time, a colleague showed that a protein found in the blood could predict which patients were at high risk for cancer. We put two and two together and realized that we could use the protein test to work out which patients would benefit from the drug.

To make things even easier for us, it turned out that the researchers who had conducted the trial had actually measured this protein in all their patients. So we wrote to them and asked whether they would share their data. They refused on the grounds that they might consider a similar analysis at some point in the future. But years have passed, no such analyses have been forthcoming and few patients are benefiting from what could be a very effective drug.

Given the enormous physical, emotional and financial toll of cancer, one might expect researchers to promote the free and open exchange of information. The patients who volunteer for cancer trials often suffer through painful procedures and harsh experimental treatments in the hope of hastening a cure. The data they provide ought to belong to all of us. Yet cancer researchers

phone call with the scientist in charge of the agenda. Only after another one-hour call with the committee itself were we allowed to submit a formal proposal—to which we received no response.

Most refusals are more blunt. "I am not prepared to release the data at this point," one researcher wrote me, even though he was a government employee and his trial, which had been published several years earlier, was federally financed.

Dr. John Kirwan, a rheumatologist from the University of Bristol in England, has studied researchers' attitudes on sharing data from clinical trials. He found that three-quarters of researchers he surveyed, as well as a major industry group, opposed making original trial data available. It is worth restating this finding: most scientists doing research on how best to help those in pain, or at risk of death, want to keep their data a secret.

Dr. Kirwan went on to ask his subjects why. Their reasons were entirely trivial: one cited the difficulty of putting together a data set (wouldn't this have to be done anyway in order to publish a paper?); another was concerned that the data might be analyzed using in-

valid methods (surely a judgment for the scientific community as a whole). This is something of a clue that the real issue here has more to do with status and career than with any loftier considerations. Scientists don't want to be scooped by their own data, or have someone else challenge their conclusions with a new analysis.

Yet this is exactly what cancer patients need. They want new results to be published as quickly as possible and to encourage a robust debate on the merits of key research findings.

An acquaintance of mine was recently diagnosed with breast cancer, and it gives me some comfort to know that there are drugs she can take that will improve her chance of cure. We know that these drugs are of benefit because more than 20 years ago, a group of Oxford statisticians persuaded researchers around the world to pool data from their breast-cancer trials.

With the rise of the Internet, sharing data has become a simple matter. Geneticists, for example, publish their raw data on a central Web site. The data from medical trials are given freely by patients. They should insist that these belong to science as a whole.



Health  
The New York Times

TUESDAY, JANUARY 22, 2008

N F5

# Typical experiences trying to obtain data from medical trials

- Needed data from the control arm of a trial to help design a study
- NIH researcher, NIH funded trial
- “I am not prepared to release the data at this point”

## Anecdote 2

- **Conducting a meta-analysis**
- **Needed proportions from a published trial that reported means and SDs**
- **“I would love to provide you with these data but my biostatistician won’t allow it”**

## Anecdote 3

- **Wanted data from a large cancer trial to illustrate a novel statistical technique**

$$\Delta_{NB} = \left( \sum_{i=1}^n x_i - p_t \right) - \left( \sum_{i=1}^n x_{i,k=1} - p_t \sum_{k=0}^1 x_k \right)$$

- **Investigators were suspicious**



**I implore you, oh great king, pity  
me, poor, little worm that I am**

- **We promised:**
  - The data would only be used for a statistical methodology study
  - We would expressly state in the paper that no clinical conclusions should be drawn
  - We would slightly corrupt the data
  - We would send a draft to the investigators and they would have veto power

# This isn't sour grapes....

- I am PI of
  - Prostate biopsy collaborative group (n=25,000)
  - Acupuncture trialists' collaboration (n=18,000)
  - Collaborations for learning curve studies in radical prostatectomy (n=15,000)

# An Empirical Evaluation of Guidelines on Prostate-specific Antigen Velocity in Prostate Cancer Detection

Andrew J. Vickers, Cathee Till, Catherine M. Tangen, Hans Lilja, Ian M. Thompson

Manuscript received March 18, 2010; revised January 3, 2011; accepted January 12, 2011.

**Correspondence to:** Andrew Vickers, PhD, Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, 307 E. 63rd St, New York, NY 10021 (e-mail: vickersa@mskcc.org).

**Background** The National Comprehensive Cancer Network and American Urological Association guidelines on early detection of prostate cancer recommend biopsy on the basis of high prostate-specific antigen (PSA) velocity, even in the absence of other indications such as an elevated PSA or a positive digital rectal exam (DRE).

**Methods** To evaluate the current guideline, we compared the area under the curve of a multivariable model for prostate cancer including age, PSA, DRE, family history, and prior biopsy, with and without PSA velocity, in 5519 men undergoing biopsy, regardless of clinical indication, in the control arm of the Prostate Cancer Prevention Trial. We also evaluated the clinical implications of using PSA velocity cut points to determine biopsy in men with low





European Association of Urology



Platinum Priority – Prostatic Disease  
*Editorial by XXX on pp. x–y of this issue*

# Individualized Estimation of the Benefit of Radical Prostatectomy from the Scandinavian Prostate Cancer Group Randomized Trial

Andrew Vickers<sup>a,\*</sup>, Caroline Bennette<sup>b</sup>, Gunnar Steineck<sup>c</sup>, Hans-Olov Adami<sup>d</sup>,  
Jan-Erik Johansson<sup>e</sup>, Anna Bill-Axelsson<sup>f</sup>, Juni Palmgren<sup>c</sup>, Hans Garmo<sup>g</sup>, Lars Holmberg<sup>h</sup>

<sup>a</sup>Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY, USA; <sup>b</sup>University of Washington, Pharmaceutical Outcomes Research and Policy Program, Seattle, WA, USA; <sup>c</sup>Karolinska Institute, Stockholm, Sweden; <sup>d</sup>Harvard School of Public Health, Boston, MA, USA; <sup>e</sup>Örebro University Hospital, Örebro, Sweden; <sup>f</sup>University Hospital Uppsala, Uppsala, Sweden; <sup>g</sup>King's College London School of Medicine, London, UK, and Regional Oncologic Centre Uppsala/Örebro, Uppsala, Sweden; <sup>h</sup>King's College London School of Medicine, London, UK, and Regional Oncologic Centre Uppsala/Örebro, Uppsala, Sweden

## Article info

### Article history:

Accepted April 5, 2012

## Abstract

**Background:** Although there is randomized evidence that radical prostatectomy improves survival, there are few data on how benefit varies by baseline risk.  
**Objective:** We aimed to create a statistical model to calculate the decrease in risk of

# Arguments against data sharing

OPEN ACCESS Freely available online



## Empirical Study of Data Sharing by Authors Publishing in PLoS Journals

Caroline J. Savage, Andrew J. Vickers\*

Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, New York, United States of America

### Abstract

**Background:** Many journals now require authors share their data with other investigators, either by depositing the data in a public repository or making it freely available upon request. These policies are explicit, but remain largely untested. We sought to determine how well authors comply with such policies by requesting data from authors who had published in one of two journals with clear data sharing policies.

**Methods and Findings:** We requested data from ten investigators who had published in either PLoS Medicine or PLoS Clinical Trials. All responses were carefully documented. In the event that we were refused data, we reminded authors of the journal's data sharing guidelines. If we did not receive a response to our initial request, a second request was made. Following the ten requests for raw data, three investigators did not respond, four authors responded and refused to share their data, two email addresses were no longer valid, and one author requested further details. A reminder of PLoS's explicit requirement that authors share data did not change the reply from the four authors who initially refused. Only one author sent an original data set.

**Conclusions:** We received only one of ten raw data sets requested. This suggests that journal policies requiring data sharing

# Arguments against data sharing 1

- **Cost and trouble of putting data set together**
  - “I would love to help you, but it would take too much time”

**It would take too long to put  
together a data set suitable for  
statistical analysis**

- **Errr.... doesn't this have to be done  
anyway?**

Top

Abstract

Background

Methods

Results

Discussion

Conclusion

Abbreviations

Competing  
interests

Authors'  
contributions

Acknowledgements

References

Pre-publication  
history

Research article

Open Access

# Acupuncture for dyspnea in advanced cancer: a randomized, placebo-controlled pilot trial [ISRCTN89462491]

Andrew J Vickers<sup>1,2</sup> ✉, Marc B Feinstein<sup>3</sup> ✉, Gary E Deng<sup>1</sup> ✉ and Barrie R Cassileth<sup>1</sup> ✉

<sup>1</sup> Integrative Medicine Service, Memorial Sloan-Kettering Cancer Center, New York, USA

<sup>2</sup> Biostatistics Service, Memorial Sloan-Kettering Cancer Center, New York, USA

<sup>3</sup> Pulmonary Service, Memorial Sloan-Kettering Cancer Center, New York, USA

✉ author email ✉ corresponding author email

*BMC Palliative Care* 2005, **4**:5 doi:10.1186/1472-684X-4-5

The electronic version of this article is the complete one and can be found online at:  
<http://www.biomedcentral.com/1472-684X/4/5>

Received: 11 April 2005

Accepted: 18 August 2005

Published: 18 August 2005

© 2005 Vickers et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Abstract**

**BMC Palliative Care**  
Volume 4

**Viewing options:**

- [Abstract](#)
- [Full text](#)
- [PDF \(251KB\)](#)
- [Additional files](#)

**Associated material:**

- [Readers' comments](#)
- [Pre-publication history](#)
- [PubMed record](#)

**Related literature:**

- [Articles citing this article on Google Scholar](#)
- [Articles citing this article on PubMed Central](#)
- [Other articles by authors on Google Scholar](#)
- [Other articles by authors on PubMed](#)
- [Related articles/pages on Google](#)
- [Related articles/pages on Google Scholar](#)
- [Related articles/pages on PubMed](#)

**Tools:**

- [Download references](#)
- [Download XML](#)
- [Email to a friend](#)
- [Order reprints](#)

## Additional files

---

### Additional File 1:

This is an Excel file with raw data from the study. The first worksheet contains the data; the second, a list of variable descriptors.

Format: XLS Size: 21KB [Download file](#)

This file can be viewed with: [Microsoft Excel Viewer](#)



Stay up to date with content from

BMC Medical Informatics  
and Decision Making

Advertise here

BMC  
Palliative Care

UNOFFICIAL  
IMPACT  
FACTOR  
1.27

Home | journals A-Z | subject areas | advanced search

search article

Acupuncture for dyspnea  
placebo-controlled pilot trial

Andrew J Vickers<sup>1,2</sup> ✉, Marc B Feinstein<sup>3</sup> ✉

<sup>1</sup>Integrative Medicine Service, Memorial Sloan-Kettering  
<sup>2</sup>Biostatistics Service, Memorial Sloan-Kettering  
<sup>3</sup>Pulmonary Service, Memorial Sloan-Kettering

✉ author email ✉ corresponding author email

BMC Palliative Care 2005, 4:5 doi:10.1186/1472-684X-4-5

Published: 18 August 2005

## Additional files

### Additional File 1:

This is an Excel file with raw data from the study. The first worksheet contains the data; the second, a list of variable descriptors.

Format: XLS Size: 21KB [Download file](#)

This file can be viewed with: [Microsoft Excel Viewer](#)

Welcome Andrew Vickers (Log off)

[Feedback](#) | [Support](#)

Open Access

BMC Palliative Care  
Volume 4

### Viewing options:

- [Abstract](#)
- [Full text](#)
- [PDF \(251KB\)](#)
- **Additional files**

### Associated material:

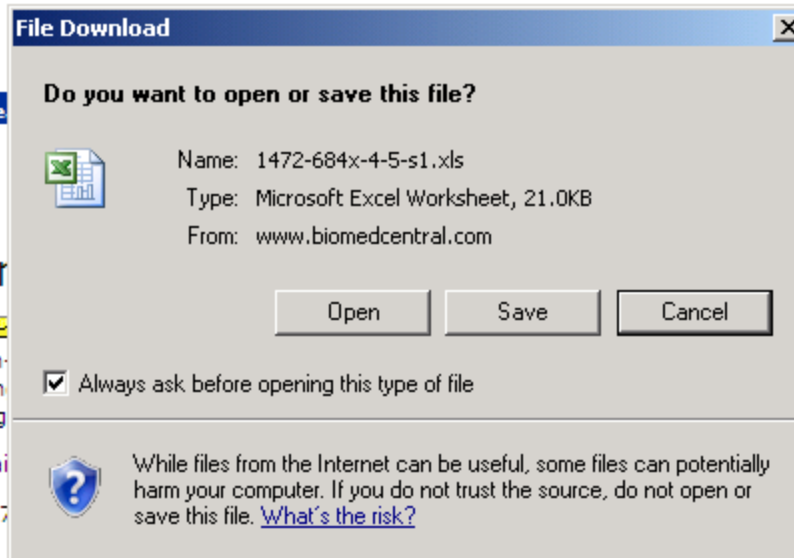
- [Readers' comments](#)
- [Pre-publication history](#)
- [PubMed record](#)

### Related literature:

- [Articles citing this article on Google Scholar on PubMed Central](#)
- [Other articles by authors on Google Scholar on PubMed](#)
- [Related articles/pages on Google on Google Scholar on PubMed](#)

### Tools:

- [Download references](#)
- [Download XML](#)
- [Email to a friend](#)
- [Order reprints](#)



# A cut and paste from the stats software

|    | A                  | B                                                   | C | D | E                                             | F | G |
|----|--------------------|-----------------------------------------------------|---|---|-----------------------------------------------|---|---|
| 1  |                    |                                                     |   |   |                                               |   |   |
| 2  | randomization      | Treatment assignment                                |   |   | A cut and paste<br>from the stats<br>software |   |   |
| 3  | group              | 1 if acupuncture, 0 for placebo                     |   |   |                                               |   |   |
| 4  | sex_code           | 1 for male, 2 for female                            |   |   |                                               |   |   |
| 5  | premean            | Mean dyspnea score, immedia                         |   |   |                                               |   |   |
| 6  | postmean           | Mean dyspnea score, immediate posttreatment period  |   |   |                                               |   |   |
| 7  | chronicpre         | Mean dyspnea score, baseline dialy diary            |   |   |                                               |   |   |
| 8  | chronicpost        | Mean dyspnea score, posttreatment dialy diary       |   |   |                                               |   |   |
| 9  | lung               | 1 if lung cancer, 0 if breast cancer                |   |   |                                               |   |   |
| 10 | breathless_at_rest | 1 if breatheless at rest, 0 otherwise               |   |   |                                               |   |   |
| 11 | age                |                                                     |   |   |                                               |   |   |
| 12 | completer          | 1 if gave follow-up data                            |   |   |                                               |   |   |
| 13 | credibility        | Credibility score                                   |   |   |                                               |   |   |
| 14 | steroidbinary      | Use of steroids (1 no 2 yes)                        |   |   |                                               |   |   |
| 15 | diuretic           | Use of diuretics (1 no 2 yes)                       |   |   |                                               |   |   |
| 16 | opiates            | Use of opiates (1 no 2 yes)                         |   |   |                                               |   |   |
| 17 | broncho            | Use of bronchodilators (1 no 2 yes)                 |   |   |                                               |   |   |
| 18 | cohort             | 1 for first prescription, 2 for second prescription |   |   |                                               |   |   |

# Data and programming code from the studies on the learning curve for radical prostatectomy

Andrew J Vickers\* and Angel M Cronin

\* Corresponding author: Andrew J Vickers [vickersa@mskcc.org](mailto:vickersa@mskcc.org)

▼ Author Affiliations

Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, NY, USA

For all author emails, please [log on](#).

*BMC Research Notes* 2010, **3**:234 doi:10.1186/1756-0500-3-234

The electronic version of this article is the complete one and can be found online at:  
<http://www.biomedcentral.com/1756-0500/3/234>

Received: 21 June 2010

Accepted: 2 September 2010

Published: 2 September 2010

© 2010 Vickers et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Abstract

# Arguments against data sharing 2

- **It might violate patient privacy**
  - “I would love to, but it is against HIPAA regulations”

| surgeon | institutionid | institution | ageatrp | cstage | psa | bg |
|---------|---------------|-------------|---------|--------|-----|----|
| ADELWSU | 7915          | WSU         | 59      | 2      | 6.8 | 8  |
| ADELWSU | 10015         | WSU         | 70      | 2      | 2   | 6  |
| ADELWSU | 9836          | WSU         | 48      | 5      | 3.5 | 6  |
| ADELWSU | 7930          | WSU         | 45      | 2      | 6.8 | 6  |
| ADELWSU | 10099         | WSU         | 63      | 2      | 6   | 7  |
| ADELWSU | 9771          | WSU         | 68      | 2      | 4.5 | 7  |
| ADELWSU | 7972          | WSU         | 60      | 2      | 6.8 | 6  |
| ADELWSU | 9710          | WSU         | 63      | 2      | 4.8 | 5  |
| ADELWSU | 9177          | WSU         | 62      | 3      | 9.3 | 7  |
| ADELWSU | 10216         | WSU         | 52      | 1      | 7   | 5  |
| AFCCF   | 2778          | CCF         | 64      | 3      | 5.4 | 6  |
| AFCCF   | 2570          | CCF         | 59      | 3      | 2.3 | 6  |
| AFCCF   | 1631          | CCF         | 56      | 2      | 7.2 | 6  |
| AFCCF   | 198           | CCF         | 64      | 3      | 7.2 | 6  |
| AFCCF   | 2601          | CCF         | 56      | 2      | 4.5 | 7  |
| AFCCF   | 2552          | CCF         | 58      | 2      | 3   | 6  |
| AFCCF   | 2612          | CCF         | 64      | 2      | 9   | 6  |

| surgeon | institutionid | institution | ageatrp | cstage | psa | bg |
|---------|---------------|-------------|---------|--------|-----|----|
| ADELWSU | 7915          | WSU         | 59      | 2      | 6.8 | 8  |
| ADELWSU | 10015         | WSU         | 70      | 2      | 2   | 6  |
| ADELWSU | 9836          | WSU         | 48      | 5      | 3.5 | 6  |
| ADELWSU | 7930          | WSU         | 45      | 2      | 6.8 | 6  |
| ADELWSU | 10099         | WSU         | 63      | 2      | 6   | 7  |
| ADELWSU | 9771          | WSU         | 68      | 2      | 4.5 | 7  |
| ADELWSU | 7972          | WSU         | 60      | 2      | 6.8 | 6  |
| ADELWSU | 9710          | WSU         | 63      | 2      | 4.8 | 5  |
| ADELWSU | 9177          | WSU         | 62      | 3      | 9.3 | 7  |
| ADELWSU | 10216         | WSU         | 52      | 1      | 7   | 5  |
| AFCCF   | 2778          | CCF         | 64      | 3      | 5.4 | 6  |
| AFCCF   | 2570          | CCF         | 59      | 3      | 2.3 | 6  |
| AFCCF   | 1631          | CCF         | 56      | 2      | 7.2 | 6  |
| AFCCF   | 198           | CCF         | 64      | 3      | 7.2 | 6  |
| AFCCF   | 2601          | CCF         | 56      | 2      | 4.5 | 7  |
| AFCCF   | 2552          | CCF         | 58      | 2      | 3   | 6  |
| AFCCF   | 2612          | CCF         | 64      | 2      | 9   | 6  |



|  |  |  | ageatrp | cstage | psa | bg |  |
|--|--|--|---------|--------|-----|----|--|
|  |  |  | 59      | 2      | 6.8 | 8  |  |
|  |  |  | 70      | 2      | 2   | 6  |  |
|  |  |  | 48      | 5      | 3.5 | 6  |  |
|  |  |  | 45      | 2      | 6.8 | 6  |  |
|  |  |  | 63      | 2      | 6   | 7  |  |
|  |  |  | 68      | 2      | 4.5 | 7  |  |
|  |  |  | 60      | 2      | 6.8 | 6  |  |
|  |  |  | 63      | 2      | 4.8 | 5  |  |
|  |  |  | 62      | 3      | 9.3 | 7  |  |
|  |  |  | 52      | 1      | 7   | 5  |  |
|  |  |  | 64      | 3      | 5.4 | 6  |  |
|  |  |  | 59      | 3      | 2.3 | 6  |  |
|  |  |  | 56      | 2      | 7.2 | 6  |  |
|  |  |  | 64      | 3      | 7.2 | 6  |  |
|  |  |  | 56      | 2      | 4.5 | 7  |  |
|  |  |  | 58      | 2      | 3   | 6  |  |
|  |  |  | 64      | 2      | 9   | 6  |  |

```

/* create anonymous ID's for patients and surgeons */
  g u1 = uniform()
  g u2 = uniform()
  sort u1 u2
  g anonpatientid = _n
  drop u1 u2
  egen anonsurgeonid = group(surgeon)

/* jitter age at surgery */
  * randomly jitter age by 1 year
  * thus ages on data set cannot be identified
  * note that age on the data set is currently given to 2 decimal places
  g anonage = int(age - 1 + uniform()*2) if uniform() < 0.33

  * put all men below age 40 to age 40
  * put all men above age 75 to age 75
  * radical prostatectomy is rare in these groups and therefore age may identify individuals
  replace anonage = 40 if anonage<40 | anonage<40
  replace anonage = 75 if anonage>75 | anonage>75

/* order the random variables next to the raw variables */
  order anonpatientid, after(uniqueid)
  order anonsurgeonid, after(surgeon)
  order anonage, after(age)

/* label the new variables */
  label var anonpatientid "anonymous patient identifier"
  label var anonsurgeonid "anonymous surgeon identifier"
  label var anonage "age at radical prostatectomy (years), with jitter of 1 year randomly added"

/* save out the data set with both identifying and non-identifying information */
  save "Files not to be sent to journal\master learning curve data identified deidentified.dta", replace

/* remove identifying information */

  * remove dates
  drop surgdate bcrdate lastfollowdate

  * remove surgeon and patient ID's, and unjittered age
  drop uniqueid institutionid surgeon age

  * remove institution
  drop institution

```

# Arguments against data sharing 3

- Other researchers might conduct invalid analyses

# Comment on a recent study of mine

- “Why .... RCTs of 100 subjects ... ? The whole thing looks like a number the authors pulled out of their nether regions and then plugged into their meta-analysis software in order to see if it would affect anything”.

**STUDY PROTOCOL**

**Open Access**

# Individual patient data meta-analysis of acupuncture for chronic pain: protocol of the Acupuncture Trialists' Collaboration

Andrew J Vickers<sup>1\*</sup>, Angel M Cronin<sup>2</sup>, Alexandra C Maschino<sup>1</sup>, George Lewith<sup>3</sup>, Hugh Macpherson<sup>4</sup>, Norbert Victor<sup>5</sup>, Karen J Sherman<sup>6</sup>, Claudia Witt<sup>7</sup>, Klaus Linde<sup>8</sup>, the Acupuncture Trialists' Collaboration

# What do you prefer?

- As a result of this analysis, we looked at publication bias and we don't think it is a problem.
- standard deviations would the difference between acupuncture and sham lose significance.

# Arguments against data sharing 4

- **Someone might delve through our data and find out that our drug doesn't work or that it has too many side-effects**

The New

NYTimes:

Go to a

DANGER

Desp  
Long

By THE  
Published:

**T**his  
wr  
Gardine  
and And

In May  
Merck,  
under si  
multibillion-dollar drug V  
made a fateful decision.

## Merck Pays

By Ed Silverman // Nov



The settlement cap  
the FDA. Although  
promotional efforts  
such as the fen-ph  
Warner-Lambert –

During the interview  
skepticism that dru  
other large drugma  
larger – to settle ch  
their drugs. Among

For Merck, the Vio  
built on emphasizing  
Now, the drugmak  
patients in the US,  
for shareholder att



# PBS NEWSHOUR

ABOUT US | LOC

Topics

Video

Recent Programs

Teacher Resources

The Rundown new

→ Online NewsHour

REGION: North America

TOPIC: Health

TRANSCRIPT

Originally Aired: April 15, 2008

Update

## Merck Faces New Questions Over Vioxx Risks, Research

Pharmaceutical company Merck knew in 2001 that its arthritis drug Vioxx could harm patients and used staffers to "ghostwrite" favorable research on the drug, articles in the Journal of the American Medical Association report. Susan Dentzer examines the charges.



THE NEWSHOUR HEALTH UNIT IS FUNDED BY A GRANT FROM  
The Robert Wood Johnson Foundation

RealAudio | Download Streaming Video

JIM LEHRER: Now, developments today on the story about what the makers of the painkilling drug Vioxx knew and when they knew it. Vioxx was pulled from the market in 2004.

Here now is our health correspondent, Susan Dentzer.

And today's revelations or today's story has to do with the report in the Journal of the American Medical Association. What was its basic finding?

SUSAN DENTZER, NewsHour health correspondent: There were actually three articles, Jim, with three different kinds of allegations, one of which was an allegation that Merck withheld from the Food and Drug Administration some data in 2001 that – the allegation is if the FDA had had this full data, it might have acted sooner against Merck and against Vioxx.



# **Data sharing as a defense against litigation**

- **We looked hard at our data**
- **We made all our data available for others to double check**

# Arguments against data sharing 5

- **Data sets contain proprietary information**

# Like what?

- In the unlikely event that there is some information that could be used by a competitor, remove that from the data set.

# Arguments against data sharing 6

- **Researchers have a right to exploit data that they may have spent years collecting**

# Why bother collecting data if you have to share it?

- You get the first paper
- Embargo periods straightforward to arrange
- Academic credit for re-used data

Commentary

## Whose data set is it anyway? Sharing raw data from randomized trials

Andrew J Vickers\*

Address: Departments of Epidemiology and Biostatistics, Medicine, Urology, Memorial Sloan-Kettering Cancer Center, NY, USA

Email: Andrew J Vickers\* - vickersa@mskcc.org

### **Appendix 2. Suggested code of conduct for analysis of published raw data**

Terminology: the "trialists" are the authors of a published report of a randomized trial; "independent investigators" are a separate group of researchers who wish to analyze the raw trial data ("new analysis")

#### ***Code of conduct for independent investigators and journals***

1. Independent investigators planning to *publish* a new analysis should contact the trialists before undertaking any analyses

# **A message to the research community, drug companies**

- **Remember open access to papers?**
- **Remember clinical trial registration?**
- **Remember inter-racial marriage?**
- **Get used to it**