Lies, Damned Lies and Statistics
Making Sense of Numbers

Nancy Reid

University of Toronto

CASLIS Toronto Chapter
February 7, 2007
collect numbers
  "statist": collection of data for the state

create rules
  "these polls are accurate to within plus or minus three percentage points 19 times out of 20"

  "We have to say that because it's a poll"

  "When the IPCC uses the phrase ‘extremely likely’, it means the scientific consensus is 95-per-cent certain"

  Jeffrey Simpson, Globe & Mail, February 3

present data (draw pictures)
Figure SPM-7. Solid lines are multi-model global averages of surface warming (relative to 1980-99) for the scenarios A2, A1B and B1, shown as continuations of the 20th century simulations. Shading denotes the plus/minus one standard deviation range of individual model annual means. The number of AOGCMs run for a given time period and scenario is indicated by the coloured numbers at the bottom part of the panel. The orange line is for the experiment where concentrations were held constant at year 2000 values. The grey bars at right indicate the best estimate of future warming.
• statistics vs statistics
• statistics as data collection vs statistical science

• collection of data
• inference from data
• presentation of data

• an information science
• an applied science
• a collaborative science
  astronomy, biology, clinical medicine, economics, finance, psychology, zoology, ...
collect data

• census
• sampling
• “New study estimating number of dead in Iraq hotly contested” (Globe and Mail, Oct 12).
• observational studies
• “Breast cancer cases soar in farm workers; Three times more likely to fall ill. Study examines agricultural link” (Toronto Star Oct 12)
• experiments
• “Long-used drug shows new promise for cancer: Therapy prescribed for metabolic disorder now found to shrink tumours in lab rats”. (Globe and Mail, Jan 17).
inference from data

• “We estimate that as of July, 2006, there have been 654 964 [392 979 – 942 636] excess Iraqi deaths as a consequence of the war”

• “The results indicate that women with breast cancer were nearly three times more likely to have worked in agriculture when compared to controls (OR = 2.8 [95% CI 1.6 – 4.8])”

• “hazard ratio was 0.91(0.83 – 1.01): this result was not statistically significant ($p = 0.07$)”

“Low-fat diet does not cut health risks”
display data
AOGCM Projections of Surface Temperatures

Global Average Surface Temperature Change (°C)
‘Politicians seduced by secrecy’

A Canadian Newspaper Association audit found that responses from government officials for information varied across the country, from outright denial of information to delays or requests for hefty fees.

**Freedom of information**

Reporters from 39 newspapers across Canada participated in the audit, making requests for government-held information of general public interest.

**Step 1. In person**

Reporters made in-person visits to public offices.

- 112 requests
- 24 records denied in full
- 16 records denied in part
- 24 records requested

**Step 2. In writing**

Reporters made 66 formal requests under relevant information laws.

- 7 records denied
- 8 records denied in part
- 10 records denied
- 7 records not responded
- 10 records not responded

**The result**

- 1 record fee requested
- 35 records denied
- 76 records released in full

**What were the requests?**

- **Federal**
  - Information from the Public Health Agency of Canada on drug stockpiles to fight pandemics.
  - 100% denied

- **Health**
  - All reports from hospitals or health districts about hospital officials or staff who received merit pay or bonuses in the last three years on record.
  - 70% denied

- **Police**
  - The total number of crimes reported and the total number of cases resolved in each community over the last five years for which data was available.
  - 40% denied

- **Pesticides**
  - From municipalities: The amount spent annually in the last three years on herbicides and pesticides and in what quantity.
  - 33% denied

* Reporter did not enter a result presumed to be denied.
Long-used drug shows new promise for cancer

Therapy prescribed for metabolic disorder now found to shrink tumours in lab rats

BY ANDRÉ PICARD
PUBLIC HEALTH REPORTER

Imagine, if you will, a drug that shrinks cancer cells and can even make tumours disappear. A couple of spoonfuls a day of powder in a glass of water is all you need.

There are no nasty side effects like nausea and hair loss, and no damage to internal organs such as with traditional chemotherapy. And it costs only about $2 a dose.

Too good to be true?

Not according to a Canadian researcher who stumbled upon the potentially new anti-cancer agent called dichloroacetate, or DCA, a drug long used to treat rare metabolic disorders.

“"This is one of the most exciting results I’ve ever had,” said Evangelos Michelakis, an associate professor of medicine at the University of Alberta in Edmonton.

"But it can’t be too enthusiastic until it works in a human being."

In a paper published today in today’s edition of the medical journal Cancer Cell, Dr. Michelakis and a group of researchers from the U of A and the University of Ottawa, report on how they were able to use DCA to shrink human lung-, breast- and brain-cancer tumours in both lab rats and in a test tube.

While this type of research in lab-
Longevity

Nobel's greatest prize

Jan 18th 2007

From The Economist print edition

Is two years extra life

NO Nobel prize has yet been awarded for the invention of an elixir of life, as far as anyone seems to be one. That, at least, is the conclusion of Matthew Rablen of the University of Warwick, in England. Dr Rablen and Dr Oswald have...
Cocaine floods the playground; [Final 3 Edition]


Jump to full text

Translate document into: [Select language]

More Like This - Find similar documents
New study estimating number of dead in Iraq hotly contested

By Estanislao Osliebowicz

More than 40,000 Iraqis have died as a result of the 2003 invasion and ensuing violence, according to a new study published by the U.S. public health experts.

But the results of the study are debated in the online version of the leading British medical journal The Lancet, an earlier finding in 2006 by researchers from the Johns Hopkins Bloomberg School of Public Health had estimated 100,000 deaths in the first 18 months of the war.

“We estimate that at least 40,000 people — 22 per cent of the population — have died in Iraq,” the authors of the Lancet study concluded. Though such death tolls might be common in times of war, the combination of low density and tests of millions of people affected has made this the deadliest international conflict of the 21st century.

U.S. President George W. Bush dismissed the report. “I don’t think the methodology is particularly well-grounded,” he said yesterday. Similarly, Iraqi government spokesman Ali-Dulaimi told Reuters: “There are targets that are exaggerated and not precise.”

Mr. Bush has previously cited the number of Iraqi deaths at 10,000. He confirmed that number yesterday.

“I stand by the figure,” he said. “One hundred thousand or whatever they preferred is... it’s not credible.”

The Lancet study is based on a survey of 3,494 households, including 18,251 individuals in 47 randomly selected sites across Iraq.

There is surprise about the size of the figure, it has more to do with our existing death tolls,” said Sarah Leah Whitson, director of the Centre for Human Rights and Human Rights Watch. “The conventional wisdom is based on shoddy information.”

The study was funded by the Massachusetts Institute of Technology.

The number of Iraqis who have died above the normal rates since March 2003 includes deaths from all causes, including those due to a rise in certain diseases and illnesses, the study said.

Nearly 29 per cent of the deaths were in the first 18 months of the war, and the majority were due to injuries from fire, the authors said.

The Lancet study is based on a survey of 3,494 households, including 18,251 individuals in 47 randomly selected sites across Iraq.

The death toll is Iraq rose in 11,5 per 1,000 people per year from 2.5 per 1,000 people per year before the invasion, they said.
Who Wants Airbags?

Evidence against Freedle's Method for Reestimating SAT Scores

A Statistician Reads the Sports Page
Is Second Place Really the First Loser?
Low-Fat Diet Does Not Cut Health Risks, Study Finds

By GINA KOLATA
Published: February 8, 2006
Everything you know about your health is wrong (again)
By MARGARET WENTE
Tuesday, February 28, 2006, Page A19

Ignore the latest study stay on a low-fat track
By LESLIE BECK
Wednesday, February 8, 2006, Page A17

Since the 1970s, reducing our fat intake has been the cornerstone of dietary advice. A high-fat diet has been associated with a greater risk of heart disease, stroke, certain cancers and obesity. But according to three reports published in the Journal of the American Medical Association, a low-fat diet does not protect postmenopausal women from breast cancers, or even heart disease.
Diet and the unborn child

The omega point

Omega-3 fatty acids are a crucial component of a healthy diet—particularly, it seems, for pregnant women wanting bright, sociable children.

There is an old joke that goes, “I’m on a seafood diet—1 see food and I eat it.” Sadly, these days, the average see-food diet doesn’t include enough seafood, even though fish are a good source of a group of nutrients known as omega-3 fatty acids.

It has been known for some time that omega-3 acids are important, but data from a long-term study of British children suggest they are even more important than had previously been realised. In particular, the amount of omega-3 in a pregnant woman’s diet helps to determine her child’s intelligence, fine-motor skills (such as the ability to manipulate small objects, and hand-eye co-ordination) and also propensity to anti-social behaviour.

That, at least, is the conclusion of Joseph Hibbeln, a researcher at America’s National Institutes of Health, who has shown that the average level of omega-3 consumed by women during pregnancy is associated with better mental performance among their children. The level of omega-3 is the highest of all in fish and seafood, and the lowest in processed foods such as bread and baked goods.

Human Nutrition.

Perhaps the most startling finding was that the children of those women who had consumed the smallest amounts of omega-3 fatty acids during their pregnancies had verbal IQs six points lower than average. That may not sound much, but it would have a serious effect on a country’s brainpower if it were widespread. And the finding is particularly pertinent because existing dietary advice to pregnant women, at least in America, is that they should limit their consumption of seafood in order to avoid exposing their fetuses to trace amounts of brain-damaging methyl mercury. Ironically, that means they avoid one of the richest sources of omega-3.

Dr Hibbeln, however, says his work shows that the benefits of eating such fish vastly outweigh the risks from the mercury in them. Indeed, in the Avon study, it was the young children of mothers with the highest intake of seafood who had the lowest levels of methyl mercury and the highest scores in tests of mental performance.

One study set out to investigate the effects of a low-mercury diet on sperm quality, which is known to be affected by mercury. The study was carried out by Dr Hibbeln and his colleagues who showed that a diet low in mercury was associated with an increased birth weight and reduced risk of low birth weight.

Studies such as this one, which rely on correlating one variable with another, are not enough to draw firm conclusions on their own, since correlation is not necessarily causation. But these results are supported by several lines of data. One is that the graphs show “dose response” curves—in other words, different levels of omega-3 have different effects. There is also a lot of experimental work showing that omega-3 have behavioural effects on adults. One of Dr Hibbeln’s other studies, for example, showed that omega-3 supplements given to violent alcoholics reduced their anger levels by a third within three months.

It also helps to have a plausible mechanism, and Dr Hibbeln thinks there is one. Research published in 2000 by a group in Canada showed that giving omega-3 supplements to piglets doubled the levels of molecules called serotonin and dopamine in the frontal cortices of the animals’ brain, which is to show...
Can chocolate save your life?
Eating just the right kind can help ward off disease, Leslie Beck writes. A11
Mortality after the 2003 invasion of Iraq: a cross-sectional cluster sample survey

Gilbert Burnham, Riyadh Lafta, Shannon Doocy, Les Roberts

Summary

Background An excess mortality of nearly 100,000 deaths was reported in Iraq for the period 2004, attributed to the invasion of Iraq. Our aim was to update this estimate.

Methods Between May and July, 2006, we did a national cross-sectional cluster sample survey. 50 clusters were randomly selected from 16 Governorates, with every cluster consisting of 40 households. Data on deaths from these households was gathered.

Findings Three misattributed clusters were excluded from the final analysis; data from 1849 households and 12,801 individuals in 47 clusters was gathered. 1474 births and 629 deaths were reported during the survey period. Pre-invasion mortality rates were 5.5 per 1000 people per year (95% CI 4.3-7.1), and 10.9-16.1 in the 40 months post-invasion. We estimate that as of June 2006, 654,965 (392,970-942,636) excess Iraqi deaths as a consequence of the war, which correspond...
three stages of sampling
first: each Governorate (province) assigned a number of clusters, for a total of 50 clusters
Baghdad (population 6.5 million) – 12 clusters; Basrah (pop 1.8 million) – 3 clusters, etc.
random selection of administrative units, with probability proportional to size
random selection of a main street within each administrative unit; random selection of a residential street crossing a main street
random selection of starting household; then continue via adjacent households until 40 have been interviewed
empty or refusing houses noted
each house asked about births, deaths, in- and out-migration
verification via death certificates
... sample survey

- 12801 household members in 1849 households
- 629 deaths since January 02; 82 pre-invasion and 547 post-invasion
- estimated pre-invasion mortality rate 5.5/1000 per year
- estimated post-invasion mortality rate 13.2/1000 per year
- excess mortality extrapolated to entire population (approx 26 million)
- leading to estimated 655,000 excess deaths
- “our estimate of excess deaths is far higher than those reported in Iraq through passive surveillance measures”
Iraqi Death Estimates Called Too High; Methods Faulted

A new estimate of the number of Iraqis who have died as a consequence of the U.S.-led invasion in March 2003 has ignited a firestorm of its own. At 400,000 to 800,000 deaths, the new number is at least 10 times higher than estimates cited by the Iraqi government and U.S.-led coalition. U.S. President George W. Bush immediately dismissed the study, characterizing its methodology as "pretty well discredited." Other Administration officials charged that the study, released with significant publicity 4 weeks before U.S. midterm elections, was politically motivated.

Researchers who spoke with Science disagree that the authors’ motives are suspect but raise several questions about the methodology of the study, which was published 11 October in The Lancet.

Experts on both sides of the debate concede that it is notoriously difficult to get an accurate count of casualties in Iraq. The Iraqi Ministry of Health has estimated up to 40,000 violent deaths so far, based on death certificates reported by hospitals and morgues. That figure falls within the range published by Iraqi Body Count, an independent London-based group opposed to the war. The statistical technique used, called cluster surveying, divides the population into different regions, neighborhoods, and households, in contrast to a random sampling of people on the streets.

The method may be sound, but several critics question the way it was carried out in this study. Madelyn Hicks, a psychiatrist and public health researcher at King’s College London in the U.K., says she “simply cannot believe” the paper’s claim that 40 consecutive houses were surveyed in a single day: “There is simply not enough time in the day,” she says, “so I have to conclude that something else is going on for at least some of these interviews.” Households may have been “prepared by someone, made ready for rapid reporting,” she says, which “raises the issue of bias being introduced.”

Lead author Gilbert Burnham, an epidemiologist at Johns Hopkins, counters that “40 adjacent households is entirely achievable in a day’s work if well organized.” Les Roberts, also at Hopkins, adds that 80% of the 547 deaths were corroborated with death certificates. The fact that hundreds of thousands of death certificates seem to have gone unregistered by the Ministry of Health is no surprise, says Roberts, because “those who have always been grossly underreported.”

Neil Johnson and Sean Gourley, physicists at Oxford University in the U.K., who have been analyzing Iraqi casualty data for a separate study, also question whether the sample is representative. The paper indicates that the survey team avoided small back alleys for safety reasons. But this could bias the data because deaths from car bombs, street-market explosions, and shootings from vehicles should be more likely on larger streets, says Johnson. Burnham counters that the team did not have access to vehicles, so that was not a factor.

They also question the accuracy of the death certificate data on which the study is based, according to John M. Robb, a public health researcher at Johns Hopkins. The certificate data are unreliable, he says, and the authors estimate that 100,000 deaths are unreported. Other critics question the way it was carried out in this study. Madelyn Hicks, a psychiatrist and public health researcher at King’s College London in the U.K., says she “simply cannot believe” the paper’s claim that 40 consecutive houses were surveyed in a single day: “There is simply not enough time in the day,” she says, “so I have to conclude that something else is going on for at least some of these interviews.” Households may have been “prepared by someone, made ready for rapid reporting,” she says, which “raises the issue of bias being introduced.”

About 8% of these extra deaths are attributed to nonviolent causes, which researchers calculate a 95% probability that the true number of violent deaths lies between 426,369 and 793,663.

Many academics spoke up in defense of the study. “I too find the survey’s estimates shocking high, … [but] the choice of method is anything but controversial,” wrote Francesco Checchi, an epidemiologist at the London School of Hygiene and Tropical Medicine on 12 October on a humanitarian Web site. The statistical technique used, called cluster surveying, divides the population into different regions, neighborhoods, and households, in contrast to a random sampling of people on the streets.

The method may be sound, but several critics question the way it was carried out in this study. Madelyn Hicks, a psychiatrist and public health researcher at King’s College London in the U.K., says she “simply cannot believe” the paper’s claim that 40 consecutive houses were surveyed in a single day: “There is simply not enough time in the day,” she says, “so I have to conclude that something else is going on for at least some of these interviews.” Households may have been “prepared by someone, made ready for rapid reporting,” she says, which “raises the issue of bias being introduced.”

Lead author Gilbert Burnham, an epidemiologist at Johns Hopkins, counters that “40 adjacent households is entirely achievable in a day’s work if well organized.” Les Roberts, also at Hopkins, adds that 80% of the 547 deaths were corroborated with death certificates. The fact that hundreds of thousands of death certificates seem to have gone unregistered by the Ministry of Health is no surprise, says Roberts, because “those have always been grossly underreported.”

Neil Johnson and Sean Gourley, physicists at Oxford University in the U.K., who have been analyzing Iraqi casualty data for a separate study, also question whether the sample is representative. The paper indicates that the survey team avoided small back alleys for safety reasons. But this could bias the data because deaths from car bombs, street-market explosions, and shootings from vehicles should be more likely on larger streets, says Johnson. Burnham counters that the team did not have access to vehicles, so that was not a factor. .
A Debate Over Iraqi Death Estimates

JOHN BOHANNON’S ARTICLE “IRAQI DEATH ESTIMATES CALLED TOO HIGH; METHODS FAULTED” (News of the Week, 20 Oct., p. 396) contains several errors that require comment.

Bohannon fails to appreciate that cluster sampling is a random sampling method. Sampling for our study was designed to give all households an equal chance of being included. In this multistage cluster sampling, random selections were made at several levels ending with the “start” house being randomly chosen. From there, the house with the nearest front door was sampled until 39 consecutive houses were selected. This usually involved a chain of houses extending into two or three adjacent streets. Using two teams of two persons each, 40 houses could be surveyed in one day. Of our 47 clusters, 13 or 28% were rural, approximating the UN estimates for the rural population of Iraq.

Bohannon states that Gilbert Burnham did not know exactly how the Iraqi team conducted its survey. The text sent to Bohannon, which he fails to cite, said, “As far as selection of the start houses, in areas where there were residential streets that did not cross the main avenues in the area selected, these were included in the random street selection process, in an effort to reduce the selection bias that more busy streets would have.” In no place does our Lancet paper say that the survey team avoided small back alleys. The methods section of the paper was modified with the suggestions of peer reviewers and the editorial staff. At no time did Burnham describe it to Bohannon as “oversimplified.”

Those who work in conflict situations know that checkpoints often scrutinize written materials carried by those stopped, and their purpose may be questioned. Unique identifiers, such as neighborhoods, streets, and houses, would pose a risk not only to those in survey locations, but also to the survey teams. Protection of human subjects is always paramount in field research. Not including unique identifiers was specified in the approval the study received from the Johns Hopkins Bloomberg School of Public Health Committee on Human Research. At no time did the teams “destroy” details, as Bohannon contends. Not recording unique identifiers does not compromise the validity of our results.

Concerning mortality estimates, Michael Spagat may be content, as Bohannon claims, with mortality data collected barely 1 year into an escalating 3.5-year war. Others might not find these so helpful.

GILBERT BURNHAM AND LES ROBERTS
Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA

Response

I DO APPRECIATE THAT CLUSTER SAMPLING relies on random samples. It is indeed the very bone of contention. “Sampling for our study was designed to give all households an equal chance of being included,” Burnham and Roberts write. But according to their methods as published in The Lancet, that is not the case.

My article reports the concerns of Sean Gourley and Neil Johnson, who point out that the starting house was always on a street “randomly selected from a list of residential streets crossing the main street.” This includes all the smaller streets, including back alleys—that do not cross a main street. This was the method used to select households and whether sample bias can explain the large number of violent deaths reported by the reported deaths.

When I asked Burnham by e-mail about this possible source of bias, he replied that “in areas where there were residential streets that did not cross the main avenues in the area selected, these were included in the random street selection process, in an effort to reduce the selection bias that more busy streets would have.” When I asked him why the published methods did not contain these details, Bohannon as “oversimplified.”

An Iraqi woman collapses after learning of the death of a relative in a bomb attack on a police car.

I stated that “the details about neighborhoods surveyed were destroyed.” The details in question are the “scraps” of paper on which streets and addresses were written to “randomly” choose households, and as Burnham and Roberts explained to me, that record has indeed been destroyed. I appreciate the difficulty of conducting a study in a combat zone and also the researchers’ desire to protect the survey team and respondents. At the same time, scientists concerned about the true number of Iraqi casualties want to know which method was used to select households and whether sample bias can explain the large number of violent deaths reported by
... sample survey

<table>
<thead>
<tr>
<th>Period</th>
<th>Mortality rates for total time post-invasion</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Invasion</td>
<td>5.5 (4.3–7.1)</td>
<td></td>
</tr>
<tr>
<td>Mar 03–Apr 04</td>
<td>7.5 (5.8–9.7)</td>
<td></td>
</tr>
<tr>
<td>May 04–May 05</td>
<td>10.9 (8.1–14.0)</td>
<td></td>
</tr>
<tr>
<td>June 05–June 06</td>
<td>19.8 (14.6–26.7)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

**Crude mortality rate**

- Non-violent mortality rate
  - 5.4 (4.1–6.8)
  - 4.5 (3.2–5.8)
  - 5.0 (3.8–6.3)
  - 6.9 (5.1–9.5)
  - 6.0 (4.8–7.5)
  - 0.523

- Violent mortality rate
  - 0.1 (0.0–0.4)
  - 3.2 (1.8–4.9)
  - 6.6 (4.0–9.8)
  - 12.0 (7.2–16.8)
  - 7.2 (5.2–9.5)
  - <0.001

- Excess mortality rate
  - 0
  - 2.6 (0.6–4.7)
  - 5.6 (2.7–8.6)
  - 14.2 (8.6–21.5)
  - 7.8 (4.7–11.2)
  - NA

Data are number of deaths per 1000 people per year (95% CI). *Post-invasion mortality rate vs pre-invasion mortality rate.

Table 3: Mortality rates by time

estimated crude mortality rate, post invasion: 13.2 [10.9 – 16.1] per 1000 per year
estimated excess number of deaths: 654 964 [392 979 – 942 636] “ accurate within these limits, 19 times out of 20”
... sample survey

- estimation of a **population quantity**
- “this poll is accurate within 3.1 percentage points, 19 times out of 20”
- “If a whole series of polls tends to confirm the same result over time, then you’ve got as close to certainty as polling can get” (Jeffrey Simpson, Globe & Mail, Feb 4)
- Iraq study #1, **March 2003 to September 2004**
- estimate of **98 000** excess deaths [8 000 – 194 000]
- new estimate **112 000** excess deaths [69 000 – 155 000]
- estimation of the “±” for standard political polls simpler than for 3 stage cluster sampling
- using the standard error to assess uncertainty exactly the same
Breaking the Embargo

Richard Horton

"Can we expect a few broken fingers?" asked one news agency journalist. "The embargo break caused me serious embarrassment," wrote the science editor for a leading international broadsheet newspaper. "There should be consequences for this sort of behavior," remarked another reporter.

These were just a few of the many angry reactions I received from journalists when three U.S. newspapers—the New York Times, Washington Post, and Wall Street Journal—broke The Lancet's news embargo on the recent Johns Hopkins survey of mortality in Iraq.

Admittedly, the story was hot. Could it be true that increased death rates after the U.S.-led invasion of Iraq in March 2003 have claimed an additional 650,000 lives? We planned to publish the research paper online. Following our standard practice (one shared with many journals), we had issued a press release and advance copy of the article to journalists a few days before the day of publication. But a dispute over who had rights to disseminate the research led journalists at those three newspapers to believe they were free to ignore our usual embargo rules. We felt stamped upon by this embargo violation. In retaliation, and at the urging of other journalists, we temporarily removed the three newspapers from our press lists (thus depriving them of advance access to papers we were about to publish).

Vincent Kiernan might, I think, be quite pleased with this outcome. In Embargoed Science, his compelling critique of the self-aggrandizing embargo system that currently rules scientist-editor-reporter relations, he stresses that journalists should not break embargoes. But he also believes that, as he bluntly puts it, "The embargo should go." What happened to The Lancet shows that the system does fracture under pressure.

Welcome news, perhaps.

In practice, an embargo means that the content of a journal is sent to journalists on the understanding that they will not write about the work until a prespecified date and time. Editors use the journal embargo as a mechanism to restrict, not promote, the communication of science to the press and public. The editors of scientific and medical journals have somehow assumed the power behind the journal embargo is a bad one: it is a mechanism to restrict, not promote, the free flow of information. It is a misimpression of immediacy. But it is a misleading impression. The governing idea behind the journal embargo is a bad one: it is a mechanism to restrict, not promote, the dissemination of research led journalists to the scientific and medical establishment.

He goes on to argue that the control of information through the embargo creates an "impression of immediacy." But it is a misleading impression. The governing idea behind the journal embargo is a bad one: it is a mechanism to restrict, not promote, the free flow of information.

But journalists like embargoes, we editors cry. It is not us who impose embargoes on news reporters. Kiernan shows that this is true, historically as well as currently. The embargo was born because journalists "demanded advance access" to scientific research. But the publishers of science soon realized what a powerful stranglehold over the press they had been given. They have exploited that power ever since.

Kiernan wins the argument about embargoes cleanly and comprehensively. Embargoes do create deference among journalists to the scientific and medical establishment. They are artificial, perpetuating the work of less-skilled journalists and giving attention to often weak and dubious science. They turn journalists away from investigating science as they would any other institution in society.

The constant stream of embargoed news releases distracts journalists from what they should be doing—namely, taking a more critical approach to their beat. The existence of this embargo-driven "pack journalism" should be antithetical to a group that usually resists any authority trying to influence what it does. It is strange that journalists acquiesce to the will of such powerful publishing organizations.
Occupation and Breast Cancer
A Canadian Case–Control Study

JAMES T. BROPHY, a,b,c MARGARET M. KEITH, a,b,c KEVIN M. GOREY, c ISAAC LUGINAAH, d ETHAN LAUKKANEN, e DEBORAH HELLYER, a ABRAHAM REINHARTZ, a ANDREW WATTERSON, b HAKAM ABU-ZAHRA, f ELEANOR MATICKA-TYNDALE, c KENNETH SCHNEIDER, f MATTHIAS BECK, g AND MICHAEL GILBERTSON b

aOccupational Health Clinics for Ontario Workers (OHCOW), Toronto, Ontario, Canada M6A 3B6
bUniversity of Stirling, Stirling UK FK9 4LA
cUniversity of Windsor, Windsor, Ontario, Canada N9B 3P4
dUniversity of Western Ontario, London, Ontario, Canada N6A 3K7
ePrince Edward Island Cancer Treatment Centre, Charlottetown, Prince Edward Island, Canada CIA 8T5
fWindsor Regional Cancer Centre, Windsor, Ontario, Canada N8W ZX3

gUniversity of York, York, UK YO10 5DD

ABSTRACT: A local collaborative process was launched in Windsor, Ontario, Canada to explore the role of occupation as a risk factor for cancer. An initial hypothesis-generating study found an increased risk for breast cancer among women aged 55 years or younger who had ever worked in farming. On the basis of this result, a 2-year case–control study was undertaken to evaluate the lifetime occupational histories of women with breast cancer. The results indicate that women with breast cancer were nearly three times more likely to have worked in agriculture when compared to the controls (OR = 2.80 [95% CI, 1.6–4.8]). The risk for those who worked in agriculture and subsequently worked in automotive-related manufacturing was further elevated (OR = 4.0 [95% CI, 1.7–9.9]). The risk for those employed in agriculture and subsequently employed in health care was also elevated (OR = 2.3 [95% CI, 1.1–4.6]). Farming tended to be among the earlier jobs worked, often during adolescence. While this article has limitations including the small sample size and the lack of information regarding specific exposures, it does provide evidence of a possible association between farming and breast cancer. The findings indicate the need for further study to determine which aspects of farming may be of biological importance and

Address for correspondence: James T. Brophy, Ph.D., 171 Kendall, Point Edward, Ontario, Canada N7V 4G6. Voice: 519-337-4627; fax: 519-337-9442. e-mail: jbrophy@ohcow.on.ca


765
... case-control study

- over 2 1/2 year period, all women with confirmed new incident primary breast cancer at Windsor Regional Cancer Center invited to participate (564/567 accepted)
- community controls chosen at random
- the controls were approximately matched by age and geographical area
- a comprehensive life history questionnaire administered (pregnancies, smoking history, family history of breast cancer, occupational history, ...)
- jobs were categorized by coders, who were blind to the case-control status of the subject
... case-control study

TABLE 2. Logistic regression-estimated odds ratios (OR) of women ever employed in agriculture, automotive-related manufacturing, or health care

<table>
<thead>
<tr>
<th></th>
<th>Odds ratio (OR)</th>
<th>95.0% C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ever worked in agriculture</td>
<td>2.8&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.6 – 4.8</td>
</tr>
<tr>
<td>Worked in agriculture and then in automotive-related manufacturing</td>
<td>4.1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.7 – 9.9</td>
</tr>
<tr>
<td>Worked in agriculture and then in health care</td>
<td>2.3&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1.1 – 4.6</td>
</tr>
<tr>
<td>Worked in automotive-related manufacturing (but never agriculture)</td>
<td>0.76</td>
<td>0.59 – 1.10</td>
</tr>
<tr>
<td>Worked in health care (but never agriculture)</td>
<td>0.85</td>
<td>0.62 – 1.17</td>
</tr>
<tr>
<td>Ever worked in retail</td>
<td>1.0</td>
<td>1.0 – 1.05</td>
</tr>
<tr>
<td>Age</td>
<td>1.0</td>
<td>1.0 – 1.02</td>
</tr>
</tbody>
</table>

<sup>a</sup>Sig. = 0.0002  
<sup>b</sup>Sig. = 0.002  
<sup>c</sup>Sig. = 0.02

estimated odds ratio: 2.8 [1.6 – 4.8]  
odds: probability ÷ (1– probability)  
ratio:  
\[
\frac{\text{prob(farm worker | cancer cases)}}{\text{prob(farm worker | controls)}} = \frac{1}{1 - \text{prob}...}
\]

no difference between the groups: odds ratio equals 1
... case-control study

- study: “women with breast cancer were nearly three times more likely to have worked in agriculture when compared to controls (OR = 2.8 [95% CI 1.6 – 4.8])”
- Toronto Star: “Women who have worked on farms are almost three times more likely to develop breast cancer than those who have never worked in agriculture”
- these can be quite different!
- we need to know the overall cancer rate in the population to compute the second odds ratio
- might be something like 1.3 times more likely; probably less
- a very common error: prosecutor’s fallacy \( \text{pr}(\text{evidence} | \text{innocent}) \); medical testing \( \text{pr}(\text{positive test } | \text{ill}) \)
“statistically significant”

- 95% CI: 95 per cent confidence interval
- estimate $\pm 2 \times$ estimated standard error
- standard error is a scale of measurement
- $p$-values provide another way to look at the same thing
- how far is the estimate away from a fixed value? (usually 0)
- on the standard error scale
- could we get such a large value if there was no difference between the groups
- the $p$-value is the estimated probability we would get such a large value if there was no difference
- small $p$-value $\Rightarrow$ data not consistent with hypothesis of no difference
- large $p$-value $\Rightarrow$ data indicates observed difference could be chance
... statistically significant

- “Cocaine floods the playground” (The Times, Mar 25 2006)
- “When the facts get in the way of a good story” (Guardian, Apr 1 2006 and www.badscience.net, Ben Goldacre)
- 1% of children sampled replied “yes” to question re cocaine use in 2004; 2% in 2005
- press release for this government survey said it found ”almost no change in patterns of drug use, drinking or smoking since 2000”
- 1% was actually 1.4%; 2% was actually 1.9%
- “What we now have is an increase of 0.5%: out of 9,000 kids, about 45 more kids saying ‘yes’ to the question.”
• Presented with a small increase like this, you have to think: is it statistically significant? Well, I did the maths, and the answer is yes, it is, in that you get a p-value of less than 0.05 ”

• “The small increase of 0.5% was only significant because it came from a large sample of 9,000 data points - like 9,000 tosses of a coin - but if they’re not independent data points, then you have to treat it, in some respects, like a smaller sample

• “Standard practice in research is to say we only accept a finding as significant if it has a p-value of 0.05 or less. But like we said, a p-value of 0.05 means that for every 100 comparisons you do, five will be positive by chance alone”
- statistically significant means, usually $p$-value smaller than 0.05 (1 in 20)
- this is computed from an estimate, and its estimated standard error
- and relies on the “bell curve” or normal distribution
- statistically significant doesn’t mean important
- how big was the observed difference? is it meaningful?
- how was the standard error estimated?
- how was the study carried out?
- how many hypotheses were tested on the same data?
Long-used drug shows new promise for cancer

Therapy prescribed for metabolic disorder now found to shrink tumours in lab rats

BY ANDRE PICARD
PUBLIC HEALTH REPORTER

Imagine, if you will, a drug that shrinks cancer cells and can even make tumours disappear.

A few spoonfuls of a day of powder in a glass of water is all you need.

There are no nasty side effects like nausea and hair loss, and no damage to internal organs such as with traditional chemotherapy.

And it costs only about $2 a dose.

Too good to be true?

Definitely not, according to a Canadian researcher who stumbled upon the potentially new anti-cancer agent called dichloroacetate, or DCA, a drug long used to treat rare metabolic disorders.

"This is one of the most exciting results I've ever had," said Evangelos Michalakis, an associate professor of medicine at the University of Alberta in Edmonton.

"But I can't be overenthusiastic until it works in a human being.

In a paper published in today's edition of the medical journal Cancer Cell, Dr. Michalakis and a group of researchers from the U of A and the University of Ottawa, report on how they were able to use DCA to shrink human lung, breast and brain cancer tumours in both lab rats and in humans.

This is the holy grail of cancer therapeutics — how to kill the cancer cells and spare normal cells.

Dr. Michalakis said DCA is "exciting" and that the research is "exciting" and that DCA has a lot of potential.

Dr. Altieri said DCA needs to move quickly from the lab into human testing. But he cautioned that there is a real possibility that will not happen.
... an experimental study

- “21 animals divided into three groups: control \( n = 5 \), DCA prevention \( n = 8 \) and DCA reversal \( n = 8 \)”
- “untreated rats rapidly developed tumours, with a constant exponential tumour growth (Figure 8A)”
- “both DCA treated groups had a significant decrease in tumour size, measured by tumour weight and maximum diameter...”
- “the decrease in tumour growth ... was associated with an increase in apoptosis (TUNEL) and a decrease in proliferation (PCNA) (Figure 8B)”
- “Statistics: values are expressed as the mean ± SEM”
... an experimental study
... an experimental study
Long-used drug shows new promise for cancer

Therapy prescribed for metabolic disorder now found to shrink tumours in lab rats

BY ANDRÉ PICARD
PUBLIC HEALTH REPORTER

Imagine, if you will, a drug that shrinks cancer cells and can even make tumours disappear. A couple of spoonfuls a day of powder in a glass of water is all you need.

There are no nasty side effects like nausea and hair loss.

This is the holy grail of cancer therapeutics — how to kill the cancer cells and spare normal cells.

EVANGELOS MICHELAKIS
... an experimental study

- “Review raises questions about drug studies on rats” (Globe and Mail, Dec 29).
- an article in the British Medical Journal comparing the results from animal experiments to the later human trials\(^1\) on six drugs used to treat a variety of conditions.
- the results from animal studies were often not borne out by human trials.
- the most serious problem identified was *publication bias*: animal studies that don’t show any effects of a certain drug don’t get published.
- review article closes with a plea for increased scientific rigour on animal experiments.
- *Economist* discussion of the same paper more measured (“Cramping tumours”, Jan 18 2007)

---

\(^1\)“Comparison of treatment effects between animal experiments and clinical trials”, Perel et al., *BMJ* online, Dec 15 2006 (doi:101136/bmj.39048.407928.BE)
Biomedical research:
Are we getting value for money?

Billions of dollars are spent on biomedical research every year, but what tangible benefits for patients result from this massive investment? Iain Chalmers suggests that the research community has become seriously dysfunctional, that patients and the public should expect better dividends from their support of biomedical research, and that the entire research community must provide better, more effective research or risk losing the support of the public.

Over the past quarter century the biomedical research community has been phenomenally successful in attracting public resources to support its activities. Money, laboratories, equipment and personnel have been provided through funding bodies supported directly by governments, through tax-exempt medical research charities, through investment in research governance infrastructure, in library subscriptions to scientific journals and so on. The public has also provided resources for biomedical research indirectly, particularly through the taxes and insurance premiums that fund health services—the ultimate source of much of the resource invested by industry in biomedical research. The public also pays for the education and training of scientists and others involved in research. Do the health dividends resulting from this massive investment represent good value for money? To put it more bluntly: what actual benefit do patients get back?

It may seem impertinent—seditious, even—to ask such a question. Obtaining reliable answers to the question undoubtedly presents substantial challenges. But, as laid bare at a recent meeting organised by the Academy of Medical Sciences, the Medical Research Council and the Wellcome Trust, there have been few systematic attempts even to address the question.

Investment in biomedical research
The UK Clinical Research Collaboration recently published a survey of government and charitable investment in biomedical research in the UK. Figure 1 shows how this investment is distributed among different kinds of research.

Over two thirds of current investment is in basic research, classified as “underpinning” and “etiological” in the survey. The remaining 30% of the investment is shared among various categories of applied research. Less than 10% is invested in treatment evaluation.

The dominance of investment in basic research reflects claims made in an influential article published in Science 30 years ago. In it, Comroe and Dripps claimed to have demonstrated that 62% of all articles judged essential for later clinical advances were the result of basic research. Unsurprisingly, the claim was widely repeated to justify bids for more resources for basic biomedical research. Although the scientific quality of the Comroe and Dripps analysis was occasionally challenged, it was not until a quarter of a century later that an attempt was made to assess whether its findings could be replicated. That assessment concluded that the Comroe and Dripps study was “not repeatable, reliable or valid.
The other headlines

small study on rats; small part of a much larger picture

comparison of lifetimes between Nobel prize winners with Nobel prize nominees

similar study for Oscar winners compared to nominees
... the other headlines

who wants airbags?:

you do, **if you wear your seatbelt too**

low-fat diet does not cut health risks

the “Rolls Royce of studies”: Women’s Health Initiative

the omega point: not yet published (diet in pregnancy/ IQ of children)
• Probably not
• case-crossover study of 10 males and 10 females with hypertension; two weeks
• yes, blood pressure went down on average
... summary

- collect, analyse and present data
- recognize and quantify variation
- are there alternative explanations
- are the comparisons valid/sensible
- do the results apply broadly or narrowly
- a picture can be worth 1000 words

- science is incremental
- statistically significant means exactly that
- studies vary enormously in quality, generalizability
- reporting of studies varies similarly
Statistics is...

- Statistics is used to help quantify uncertainty
- Statistics is part of science
- Statistics is not all “damned lies”!
- Statistics is actually fun
“But don’t you think that life with a statistician would be rather, shall we say, humdrum?”

Lamia was silent. She felt reluctant to discuss with her aunt the surprising depth of emotional possibility which she had discovered below Edward’s numerical veneer.

“It’s not the figures themselves,” she said finally, “It’s what you do with them that matters.”

Further reading

- *Significance*, published quarterly, Royal Statistical Society
- [www.badscience.net](http://www.badscience.net) Ben Goldacre
- *Chance* magazine, published quarterly, American Statistical Association
Sources


