

# De journalistieke toets der kritiek

- Meer ongelukken bij mooi weer
- Tranen dempen mannenlust
- Nooit meer rood vlees
- Gemiste penalty kost levens



## Meer ongelukken bij mooi weer

Door onze redactie wetenschap ROTTERDAM, 17 NOV. Bij mooi weer belanden er meer mensen op de spoedeisende hulp van het ziekenhuis dan bij slecht weer. En bij volle maan zijn het er minder dan bij nieuwe maan. Dat schrijven onderzoekers van het UMC Groningen deze week in The Journal of Trauma. Zij koppelden gegevens van ruim 350.000 mensen die in de afgelopen 36 jaar de Groningse eerste hulp bezochten, aan de dagelijkse gegevens van het KNMI voor het vliegveld Eelde. De onderzoekers vermoeden dat mensen bij mooi weer vaker buiten actief zijn en daarmee meer risico lopen op ongevallen. Bij volle maan zouden er 's nachts minder ongelukken gebeuren dankzij het extra licht.

NRC Handelsblad 17 november 2009

# journalistieke toets

De onderzoekers vonden 'ruim 350 000' ongevallen in 36 jaar. Kan dat ongeveer kloppen?

#### ORIGINAL ARTICLE

Relation of the Weather and the Lunar Cycle With the Incidence of Trauma in the Groningen Region Over a 36-Year Period

Wouter Stomp, MD, Vaclav Fidler, PhD, Henk-Jan ten Duis, MD, PhD, and Maarten W. N. Nijsten, MD, PhD

## exclusion of special days. The mean number of daily admissions was 26.9 (SD 8.2) with a range of 2 to 78. The mean age

Methods: We extracted the daily number of trauma patients treated at the emergency department over 36 years (1970–2005) from the trauma database of our regional hospital. For each patient, age, sex, cause of injury, and severity of injury were recorded. This was combined with daily meteorological data including temperature, precipitation, sunshine, humidity, air pressure, and wind as well as the lunar phase. We also related the rate of change of these parameters with the incidence of injuries. A qualitative weather variable derived from temperature, sunshine duration, and precipitation was defined as bad, normal, or good. Periodicities were adjusted for with Poisson regression spline fitting analysis.

Results: Several weather variables were related with the number of injuries. For most of these, better weather conditions were associated with an increase in trauma incidence. Good weather, which was present on 16.5% of the days, resulted in 10.1% (9.3–11.4 95% CI) more traumas compared with normal weather. Full moon was associated with a 2.1% (1.1–3.0 95% CI) lower trauma incidence than new moon.

ment. For this purpose, all patients referred to our regional trauma service over a period spanning several decades were studied. We hypothesized that aside from obvious patterns related to the calendar, the quality of weather might have a significant impact on the number of trauma patients.

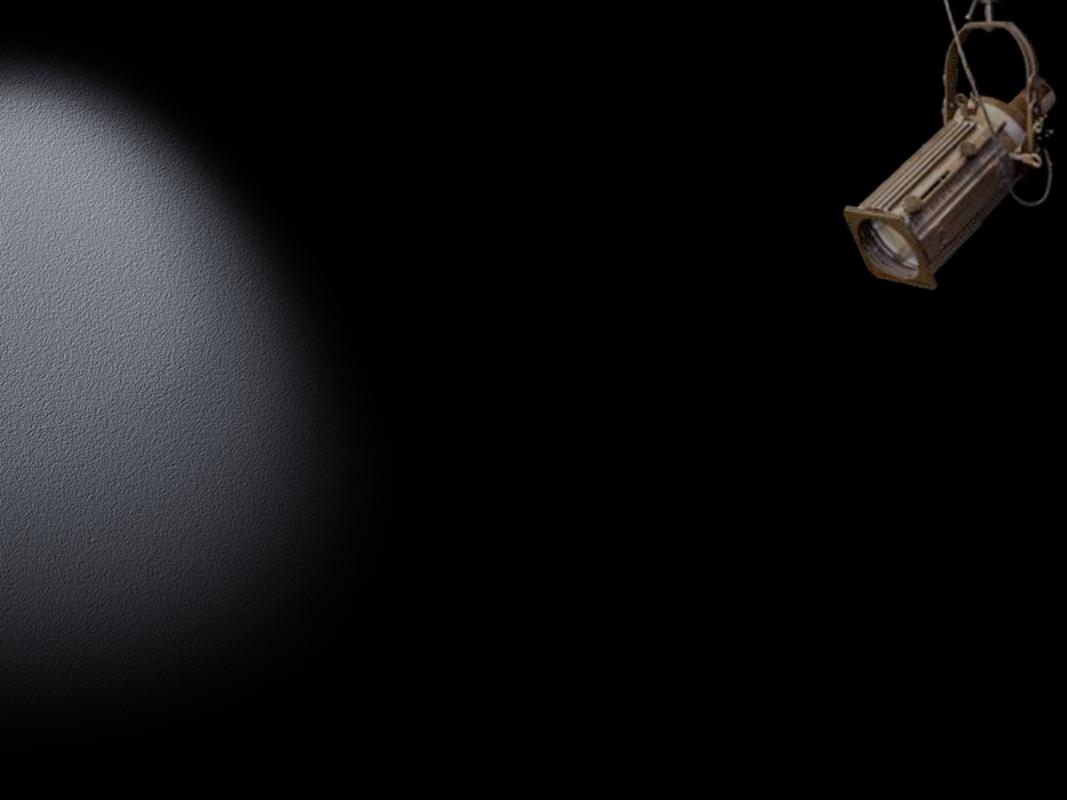
#### **PATIENTS AND METHODS**

#### **Patients**

The University Medical Center Groningen is situated in the northern part of the Netherlands. It is a Level I trauma center and is 1 of the 10 recognized trauma centers in the Netherlands and 1 of the 4 trauma centers that have a helicopter service available. Our catchment area covers the Northern part of the Netherlands with a population of approx-

# De journalistieke toets

En hoeveel procent meer zou 'meer' mensen op de spoedeisende hulp zijn?



#### ORIGINAL ARTICLE

Relation of the Weather and the Lunar Cycle With the Incidence of Trauma in the Groningen Region Over a 36-Year Period

in trauma incidence. Good weather, which was present on 16.5% of the days, resulted in 10.1% (9.3–11.4 95% CI) more traumas compared with normal weather. Full moon was associated with a 2.1% (1.1–3.0 95% CI) lower trauma incidence than new moon.

ical data including temperature, precipitation, sunshine, humidity, air pressure, and wind as well as the lunar phase. We also related the rate of change of these parameters with the incidence of injuries. A qualitative weather variable derived from temperature, sunshine duration, and precipitation was defined as bad, normal, or good. Periodicities were adjusted for with Poisson regression spline fitting analysis.

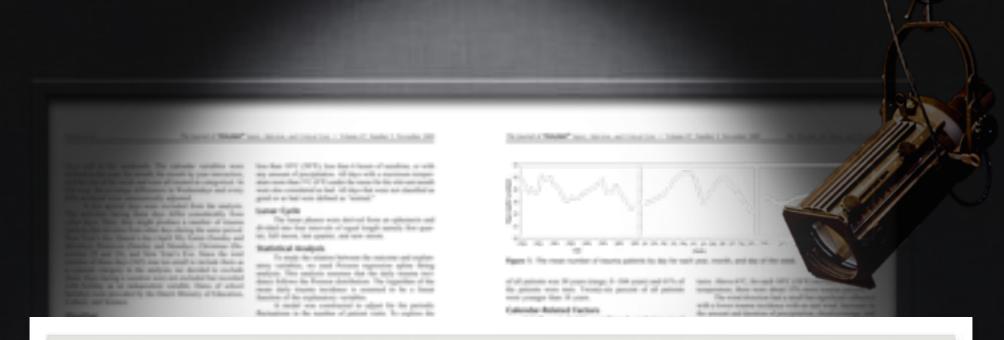
Results: Several weather variables were related with the number of injuries. For most of these, better weather conditions were associated with an increase in trauma incidence. Good weather, which was present on 16.5% of the days, resulted in 10.1% (9.3–11.4 95% CI) more traumas compared with normal weather. Full moon was associated with a 2.1% (1.1–3.0 95% CI) lower trauma incidence than new moon.

significant impact on the number of trauma patients.

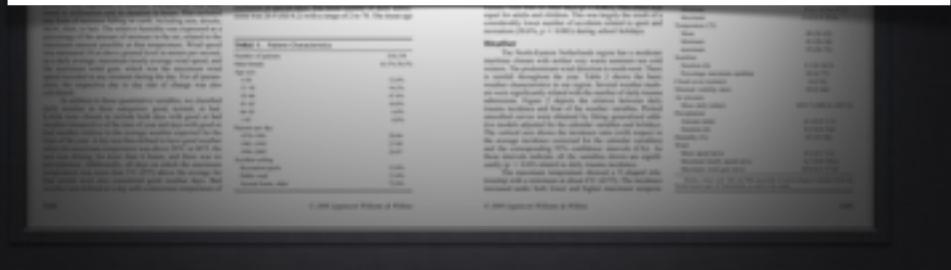
#### PATIENTS AND METHODS

#### **Patients**

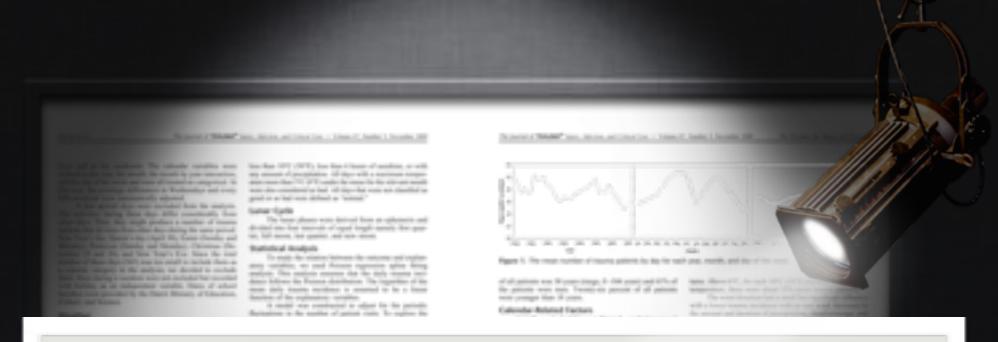
The University Medical Center Groningen is situated in the northern part of the Netherlands. It is a Level I trauma center and is 1 of the 10 recognized trauma centers in the Netherlands and 1 of the 4 trauma centers that have a helicopter service available. Our catchment area covers the Northern part of the Netherlands with a population of approx-



The wind direction had a small but significant influence with a lower trauma incidence with an east wind. Increases in



Journal of Trauma, november 2009



The wind direction had a small but significant influence with a lower trauma incidence with an east wind. Increases in the amount and duration of precipitation, cloud coverage, and relative humidity were associated with a slight decrease in the number of patients.

Relation of the Weather and the Lunar Cycle With the Incidence of Trauma in the Groningen Region Over a 36-Year Period: Erratum

n the article that appeared on pages 1103–1108, in volume 67, number 5, the authors stated that there was a lower trauma incidence with east wind. This should have been a lower trauma incidence with west wind.

#### **Accuracy of Data in Abstracts** of Published Research Articles

Roy M. Pitkin, MD

Mary Ann Branagan

Leon F. Burmeister, PhD

HE ABSTRACT ACCOMPANYING A research article, because it is often the only part of the article that will be read, should reflect fully and accurately the work reported. We observed in 1 medical specialty journal that **Context** The section of a research article most likely to be read is the abstract, and therefore it is particularly important that the abstract reflect the article faithfully.

Objective To assess abstracts accompanying research articles published in 6 medical journals with respect to whether data in the abstract could be verified in the article

Design Analysis of simple random samples of 44 articles and their accompanying abstracts published during 1 year (July 1, 1996-June 30, 1997) in each of 5 major general medical journals (Annals of Internal Medicine, BMJ, JAMA, Lancet, and New England Journal of Medicine) and a consecutive sample of 44 articles published during 15 months (July 1, 1996-August 15, 1997) in the CMAJ.

Main Outcome Measure Abstracts were considered deficient if they contained data

**Conclusions** Data in the abstract that are inconsistent with or absent from the article's body are common, even in large-circulation general medical journals.

Journal of the American Medical Association 24 maart 1999

Articles studied included simple random was .05 and power was 0.8, yielding a samples of reports of original research (including meta-analyses but not other types of reviews) appearing in 5 medical journals between July 1, 1996, and June 30, 1997 (Annals of Internal Medicine, BMJ, JAMA, Lancet, and New England Journal of Medicine); all articles appearing in a sixth journal CMAJ (Canadian Medical Associa-August 15, 1997, were also studied. Ad-

For editorial comment see p 1129.

consecutive cohort of all 44 articles pubtion Journal), between July 1, 1996, and lished from July 1, 1996, through August 15, 1997. For each selected article, the abstract ticle was accompanied by an abstract and was scrutinized by 1 of 3 examiners who

published more than 44 research articles in the 2 volumes studied (July 1,

1996-June 30, 1997), we selected a com-

puter-generated simple random sample

sought to relate it to its source in the body of the article, including tables and figures. Two types of discrepancies were sought: (1) data given differently in the abstract and the body and (2) data given differently in the abstract and the body and (2) data given differently in the abstract and the body and (2) data given differently in the abstract and the body and (2) data given differently in the abstract and the body and (2) data given differently in the abstract and the body and (2) data given differently in the abstract and the body and (2) data given differently in the abstract and the body and (2) data given differently in the abstract and the body and (2) data given differently in the abstract and the body and (2) data given differently in the abstract and the body and (3) data given differently in the abstract and the body and (3) data given differently in the abstract and the body and (3) data given differently in the abstract and the body and (3) data given differently in the abstract and the body and (3) data given differently in the abstract and the body and (3) data given differently in the abstract and the body and (4) data given differently in the abstract and the body and (3) data given differently in the abstract and the body and (4) data given differently in the abstract and the body and (5) data given differently in the abstract and the body and (5) data given differently in the abstract and the body and (5) data given differently in the abstract and the body and (6) data given differently in the abstract and the body and (6) data given differently in the abstract and the body and (6) data given differently in the abstract and the body and (6) data given differently in the abstract and the body and (6) data given differently in the abstract and the body and (6) data given differently in the abstract and the body and (6) data given differently in the abstract and the body and (6) data given differently in the abstract and the body and (6) data given differently in the abstract and the bod

ing was done appropriately, and the projected sample size of 44 from each rounded value appeared in the abstract journal. From each of the 5 journals that and the more detailed value in the body.

ing deficiencies were compared across journals by  $\chi^2$  analysis. On the basis of normal approximation, 95% confiof 44. From the CMAJ, we analyzed a dence intervals (CIs) were calculated for each proportion. We also performed a 44 articles and having these examined by



## Misrepresentation of Randomized Controlled Trials in Press Releases and News Coverage: A Cohort Study

Amélie Yavchitz<sup>1,2,3</sup>, Isabelle Boutron<sup>1,2,3</sup>\*, Aida Bafeta<sup>1,2,3</sup>, Ibrahim Marroun<sup>4</sup>, Pierre Charles<sup>4</sup>, Jean Mantz<sup>5</sup>, Philippe Ravaud<sup>1,2,3</sup>

1 INSERM, U738, Paris, France, 2 Centre d'Epidémiologie Clinique, AP-HP (Assistance Publique des Hôpitaux de Paris), Hôpital Hôtel Dieu, Paris, France, 3 Université Paris Descartes, Sorbonne Paris Cité, Faculté de Médecine, Paris, France, 4 Department of Internal Medicine, Hôpital Foch, Suresnes, France, 5 Department of Anesthesiology and Critical Care, Beaujon University Hospital, Clichy, France

#### Abstract

Background: Previous studies indicate that in published reports, trial results can be distorted by the use of "spin" (specific reporting strategies, intentional or unintentional, emphasizing the beneficial effect of the experimental treatment). We aimed to (1) evaluate the presence of "spin" in press releases and associated media coverage; and (2) evaluate whether findings of randomized controlled trials (RCTs) based on press releases and media coverage are misinterpreted.

Methods and Findings: We systematically searched for all press releases indexed in the EurekAlert! database between December 2009 and March 2010. Of the 498 press releases retrieved and screened, we included press releases for all two-arm, parallel-group RCTs (n=70). We obtained a copy of the scientific article to which the press release related and we systematically searched for related news items using Lexis Nexis. "Spin," defined as specific reporting strategies (intentional or unintentional) emphasizing the beneficial effect of the experimental treatment, was identified in 28 (40%) scientific article abstract conclusions and in 33 (47%) press releases. From bivariate and multivariable analysis assessing the journal type, funding source, sample size, type of treatment (drug or other), results of the primary outcomes (all nonstatistically significant versus other), author of the press release, and the presence of "spin" in the abstract conclusion, the only factor associated, with "spin" in the press release was "spin" in the article abstract conclusions (relative risk [RR] 5.6, [95% CI 2.8–11.1], p<0.001). Findings of RCTs based on press releases were overestimated for 19 (27%) reports. News Items were identified for 41 RCTs; 21 (51%) were reported with "spin," mainly the same type of "spin" as those identified in the press release and article abstract conclusion. Findings of RCTs based on the news item was overestimated for ten (24%) reports.

Conclusion: "Spin" was identified in about half of press releases and media coverage. In multivariable analysis, the main factor associated with "spin" in press releases was the presence of "spin" in the article abstract conclusion.

Please see later in the article for the Editors' Summary.

Citation: Yavchitz A, Boutron I, Bafeta A, Marroun I, Charles P, et al. (2012) Misrepresentation of Randomized Controlled Trials in Press Releases and News

"Spin" was identified in about half of press releases and media coverage. In multivariable analysis, the main factor associated with "spin" in press releases was the presence of "spin" in the article abstract conclusion.

PLoS Medicine september 2012



Isliidan traact an calco

sualiteit En een over het ten shak de Tilbu hoets, d doet, de vindt o onderzo sentrane De ontd zoekers Weizma tranen ze die t hen roo

mannelijke lust. Want er zitten feromonen in tranen. Dat bleek donderdag uit een artikel in *Science*.

Veel wetenschappers bogen zich al

over de raadselachtige mensentranen, maar seksualiteit was in dat onderzoek nooit belangrijk.

Maar intussen hadden die mannen wel minder zin in seks gekregen, bleek in een reeks experimenten die in totaal drie jaar in beslag nam. Ze vonden vrouwengezichten minder seksueel aantrekkelijk. En van een verdrietige film, we meestal van zeer dichtbij aan tranen blootgesteld'



#### **Human Tears Contain a Chemosignal**

Shani Gelstein,  $^{1_*}$  Yaara Yeshurun,  $^{1_*}$  Liron Rozenkrantz,  $^1$  Sagit Shushan,  $^{1,2}$  Idan Frumin,  $^1$  Yehudah Roth,  $^2$  Noam Sobel  $^1$ †

Emotional tearing is a poorly understood behavior that is considered uniquely human. In mice, tears serve as a chemosignal. We therefore hypothesized that human tears may similarly serve a chemosignaling function. We found that merely sniffing negative-emotion—related odorless tears obtained from women donors induced reductions in sexual appeal attributed by men to pictures of women's faces. Moreover, after sniffing such tears, men experienced reduced self-rated sexual arousal, reduced physiological measures of arousal, and reduced levels of testosterone. Finally, functional magnetic resonance imaging revealed that sniffing women's tears selectively reduced activity in brain substrates of sexual arousal in men.

harles Darwin suggested that expressive behaviors initially served emotion-relevant functions, before evolving to serve as emotion-signals alone (1, 2). Thus, the behavior of emotional tearing, considered uniquely human

<sup>1</sup>Department of Neurobiology, Weizmann Institute of Science, Rehovot 76100, Israel. <sup>2</sup>Department of OtoRhinoLaryngology and Head and Neck Surgery, Edith Wolfson Medical Center, Holon 58100, Israel.

\*These authors contributed equally to this work. †To whom correspondence should be addressed. E-mail: noam.sobel@weizmann.ac.il (3), is a paradox: Whereas tears clearly serve as an emotional signal (4), tears were not related to any emotionally relevant function. Despite psychological theories on the meaning of tears (5, 6) and biological theories describing tears as an adaptation related to their eye-protective nature (3) or a mechanism for expelling toxic substances (7), the functional significance of emotional tears remains unknown (8).

Tears are drops of liquid produced by the lacrimal, accessory lacrimal, and Meibomian glands, which contain proteins, enzymes, lipids, metabo-

Science, 8 januari 2011

ARY 2011 VOL 331 **SCIENCE** www.sciencemag.org





2.29, P < 0.03] (Fig. 3, E and F). Finally, and critically, levels of salivary testosterone were progressively lower after sniffing tears as compared to the baseline period [baseline testosterone =  $151.96 \pm 76$  pg/ml, last testosterone =  $132.66 \pm$ 63.1 pg/ml, t(49) = 3.3, P < 0.001] (Fig. 3G), an effect not evident for saline [baseline testosterone =  $154.8 \pm 74.4 \text{ pg/ml}$ , last tes 0.42 ne =  $154.34 \pm$ 101.8 pg/ml, t(49) = 0.81, P = 0.96 (Fig. 3H). Reductions in testosterone are a significant indicator of reductions in sexual arousal in men (21).

## **BMC Medical Research Methodology**



Research article

Open Access

Incongruence between test statistics and P values in medical papers Emili García-Berthou\* and Carles Alcaraz

Address: Department of Environmental Sciences, University of Girona, E-17071 Girona, Spain

Email: Emili García-Berthou\* - emili.garcia@udg.es; Carles Alcaraz - carles.alcaraz@udg.es

\* Corresponding author

Published: 28 May 2004

Received: 02 January 2004 Accepted: 28 May 2004

BMC Medical Research Methodology 2004, 4:13

This article is available from: http://www.biomedcentral.com/1471-2288/4/13

© 2004 García-Berthou and Alcaraz; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

## **Results:** 11.6% and 11.1% of the statistical results published in *Nature* and *BMJ* respectively during 2001 were incongruent, probably mostly due

BMJ, respectively. In 12% of the cases, the significance level might change one or more orders of magnitude. The frequencies of the last digit of statistics deviated from the uniform distribution and suggested digit preference in rounding and reporting.

**Conclusions:** This incongruence of test statistics and *P* values is another example that statistical practice is generally poor, even in the most renowned scientific journals, and that quality of papers should be more controlled and valued.

#### **Background**

Statistics is a difficult topic to teach and learn and there is ample evidence that its application is often faulty in medicine [1-6] as well as in many other scientific disciplines. Errors include aspects of design, analysis, and reporting and interpretation. Although there has recently been considerable effort to improve and standardise the reporting of medical research (e.g., the CONSORT statement for randomised controlled trials [7]), there is almost no literature demonstrating the incorrect computation or reporting of results beyond general deficiencies of computer packages [8,9] or some well-scrutinized data such as Benford's original data [10]. Beyond deficiencies of software

such numerical errors may later originate in the transcription of results from computer outputs to reports and manuscripts, wrong rounding of results, or uncorrected typesetting errors. We investigated this question by checking the statistical results reported in all the papers of volumes 409–412 of *Nature* (2001) and some papers in vol. 322–323 of *BMJ* (2001). We show that the occurrence of errors is very high and we review ways to improve current practice.

#### Methods

Given an observed test statistic and its degrees of freedom (df), one may compute the observed *P* value o

BMC Med Res Methodol. 28 mei 2004



Voor de film

Na de film

Tranen

I52

133

Zout

I55

154

Gemiddelde testosteronspiegels van 50 mannen voor en na het zien van een treurige film



Zou dit verschil statistisch significant zijn?

#### **Erroneous** analyses of interactions in neuroscience: a problem of significance

Sander Nieuwenhuis<sup>1,2</sup>, Birte U Forstmann<sup>3</sup> & Eric-Jan Wagenmakers<sup>3</sup>

In theory, a comparison of two experimental effects requires a

statistical test on their difference. In practice, this comparison

is often based on an incorrect procedure involving two separate tests in which researchers conclude that effects differ when one effect is significant (P < 0.05) but the other is not (P > 0.05). We reviewed 513 behavioral, systems and cognitive neuroscience articles in five top-ranking journals (Science, Nature, Nature Neuroscience, Neuron and The Journal of Neuroscience) and found that 78 used the correct procedure and 79 used the incorrect procedure. An additional analysis suggests that incorrect analyses of interactions are even more common in cellular and molecular neuroscience. We discuss scenarios in which the erroneous procedure is

That is, as famously noted by Rosnow and Rosenthal<sup>2</sup>, "surely, God loves the 0.06 nearly as much as the 0.05". Thus, when making a comparison between two effects, researchers should report the statistical significance of their difference rather than the difference between their significance levels.

Our impression was that this error of comparing significance levels is widespread in the neuroscience literature, but until now there were no aggregate data to support this impression. We therefore examined all of the behavioral, systems and cognitive neuroscience studies published in four prestigious journals (Nature, Science, Nature Neuroscience and Neuron) in 2009 and 2010 and in every fourth issue of the 2009 and 2010 volumes of The Journal of Neuroscience. In 157 of these 513 articles (31%), the authors describe at least one situation in which they might be tempted to make the error. In 50% of they

(Science, Nature, Nature Neuroscience, Neuron and The Journal of Neuroscience) and found that 78 used the correct procedure and 79 used the incorrect procedure. An additional

> between significant and not significant need not itself be statistically an enormous difference between the two effect sizes or because the significant<sup>1</sup>. Consider an extreme scenario in which traininginduced activity barely reaches significance in mutant mice (for example, P = 0.049) and barely fails to reach significance for control the error cases, the error may have had serious consequences. In all mice (for example, P = 0.051). Despite the fact that these two P values of these cases, the nonsignificant difference, although smaller in size, lie on opposite sides of 0.05, one cannot conclude that the training was in the same direction as the significant difference. In addition,

Institute for Brain and Cognition, Leiden, The Netherlands. <sup>3</sup>Cognitive Science Center Amsterdam, University of Amsterdam, Amsterdam, The Netherlands. Correspondence should be addressed to S.N. (s.nieuwenhuis@fsw.leidenuiv.nl).

reported methodological information allowed us to determine the approximate significance level. Nonetheless, in roughly two thirds of effect for mutant mice differs statistically from that for control mice. the methodological information did not allow us to determine the significance level of the missing interaction test. We have no way of assessing the severity of these cases. Most of the errors may not have severe implications. In some cases, however, the error may contribute

Because of our background expertise, our main analysis focused

Nature Neuroscience september 2011

#### **Erroneous** analyses of interactions in neuroscience: a problem of significance

Sander Nieuwenhuis<sup>1,2</sup>, Birte U Forstmann<sup>3</sup> & Eric-Jan Wagenmakers<sup>3</sup>

In theory, a comparison of two experimental effects requires a statistical test on their difference. In practice, this comparison is often based on an incorrect procedure involving two separate tests in which researchers conclude that effects differ when one effect is significant (P < 0.05) but the other is not (P > 0.05). We reviewed 513 behavioral, systems and cognitive neuroscience articles in five top-ranking journals (Science, Nature, Nature Neuroscience, Neuron and The Journal of Neuroscience) and found that 78 used the correct procedure and 79 used the incorrect procedure. An additional analysis suggests that incorrect analyses of interactions are even more common in cellular and molecular neuroscience. We discuss scenarios in which the erroneous procedure is

That is, as famously noted by Rosnow and Rosenthal<sup>2</sup>, "surely, God loves the 0.06 nearly as much as the 0.05". Thus, when making a comparison between two effects, researchers should report the statistical significance of their difference rather than the difference between their significance levels.

Our impression was that this error of comparing significance levels is widespread in the neuroscience literature, but until now there were no aggregate data to support this impression. We therefore examined all of the behavioral, systems and cognitive neuroscience studies published in four prestigious journals (Nature, Science, Nature Neuroscience and Neuron) in 2009 and 2010 and in every fourth issue of the 2009 and 2010 volumes of The Journal of Neuroscience. In 157 of these 513 articles (31%), the authors describe at least one situation in which they might be tempted to make the error. In 50% of thes.

in Nature Neuroscience in 2009 and 2010 (the first five Articles in each issue). We did not find a single study that used the correct statistical procedure to compare effect sizes. In contrast, we found at least 25

> between significant and not significant need not itself be statistically an enormous difference between the two effect sizes or because the significant<sup>1</sup>. Consider an extreme scenario in which traininginduced activity barely reaches significance in mutant mice (for example, P = 0.049) and barely fails to reach significance for control the error cases, the error may have had serious consequences. In all mice (for example, P = 0.051). Despite the fact that these two P values of these cases, the nonsignificant difference, although smaller in size, lie on opposite sides of 0.05, one cannot conclude that the training was in the same direction as the significant difference. In addition,

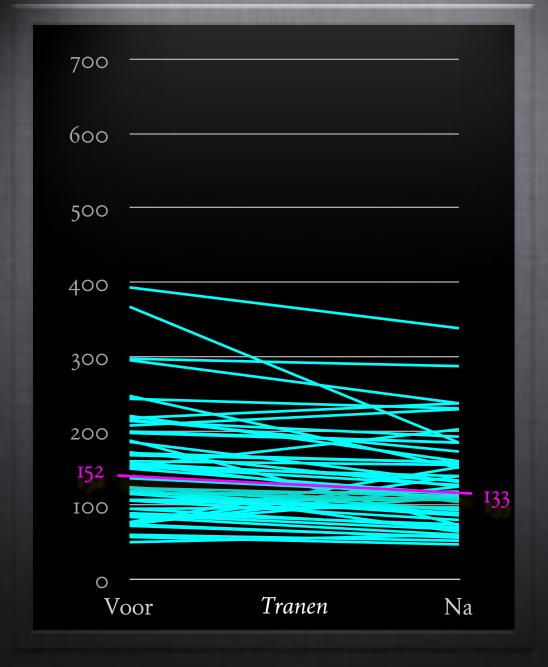
Institute for Brain and Cognition, Leiden, The Netherlands. <sup>3</sup>Cognitive Science Center Amsterdam, University of Amsterdam, Amsterdam, The Netherlands. Correspondence should be addressed to S.N. (s.nieuwenhuis@fsw.leidenuiv.nl).

reported methodological information allowed us to determine the approximate significance level. Nonetheless, in roughly two thirds of effect for mutant mice differs statistically from that for control mice. the methodological information did not allow us to determine the significance level of the missing interaction test. We have no way of assessing the severity of these cases. Most of the errors may not have severe implications. In some cases, however, the error may contribute

Because of our background expertise, our main analysis focused

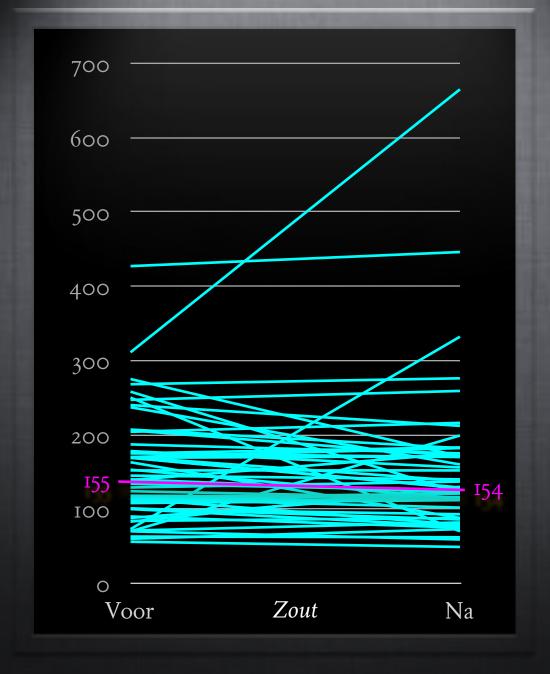


	Voor de film	Na de film
Trane n	152	133
Zout	155	154





	Voor de film	Na de film
Trane n	152	133
Zout	155	154



## edelijden, troost en seks

Door vrouwensmart vergaat mannen de zin in seks. Hester van Santen

sualite En een over h

ten sha de Til hoets, doet,

vindt sentra De on

zoeke Weizi trane over het hoofd gezien. "Ik heb twee dagen zitten shaken toen ik het stuk las", zo vreemd vond de Tilburgse psychologiehoogleraar Ad Vingerhoets, die al twintig jaar onderzoek naar huilen

Maar intussen hadden die mannen wel minder zin in seks gekregen, bleek in een reeks experimenten die in totaal drie jaar in beslag nam. Ze vonden vrouwengezichten minder seksueel

aantrekkeliik. En van een verdrietige film,

we meestal van zeer dichtbij aan tranen blootgesteld'



contributes substantially to premature death," said senior author Frank Hu, professor

RQUIN SOMEWHARM OTHER YEAR, IMPROVING PROCESSES

A new study by Harvard Subsol of Public Health (HSPH) researchers has found that red must consumption is associated with an increased risk of total, cardiovascular, and cancer mortality. The results also showed that substituting other hoolthy protein sources, such as

fish, poultry, muts, and legumes, was associated with a lower risk of

restricts and epidemiology of the Harvard Science of Public Faults.

resk) sortifices substantivly is premature static," soil sense suffer Frank tru, professor of



### 'Rood viees verhoogt kans op vroege dood'





WWW. M Har 2012 25:40-58: +0050



Iedere dag biefstuk, rosbief of worst is zeer ongezond. Mensen die veel rood vlees eten, verlagen hun levensverwachting aanzienlijk.

and Medical School

and views: sen oxine portir orverworks read verfreegt het risite op overlijden met maar

is not maximum, on solare on worst sign

#### Slackt your hat had

and other branes.

Westerland toget 10 propert report term up didulifies hart- at eastablies. Dehave dat (i) harder intigers is sold for precent hoper.

Under road vises wanth vises verstaan dat afficienting is van look, verben en schaep. Rip, kalkoen en ander plukryvee vallen er buten. Yooral versedigde wetten en zout waken vices en viseswaren angazend.

#### Ext.met.mete

Als viscounters stick socialish beganises talt 42 great visco per day of 66n blefstak. per week minder, dan zoudon se daarmoe 10 dagen extra leveratijd kunnen wished.

#### LEES DOW

Of recordo plants raken as

New American Ser (Alden)

Tanderberold-test: we'ke is het beste?

#### MEER COITE NIEUWS

80 tire; 3005	Your restanded buby, video past do would over
0.0 km 4004	Mysterious weden geffind in ecoson (yelles)
B ther Jord	"Morthonen kebben Dranges gedospt"
F 110 2017	16 dit schiderg 87 miljoen waar81
9 mm 2002	Bilde Navy Seats helden Barvetin Aut of Valor
9 For 2017	Zarvekinsmender werdt pag (video)
Brief 2003	"De walf must bring"
9 mar 2007	De verkoetende 'saperbelse', hot er red?'
F ## 250	With worth jour reteming-connect?
0 mm 2007	Tireplach needween ap konst











#### **ONLINE FIRST**

#### **Red Meat Consumption and Mortality**

#### Results From 2 Prospective Cohort Studies

An Pan, PhD; Qi Sun, MD, ScD; Adam M. Bernstein, MD, ScD; Matthias B. Schulze, DrPH; JoAnn E. Manson, MD, DrPH; Meir J. Stampfer, MD, DrPH; Walter C. Willett, MD, DrPH; Frank B, Hu, MD, PhD

**Background:** Red meat consumption has been associated with an increased risk of chronic diseases. However, its relationship with mortality remains uncertain.

**Methods:** We prospectively observed 37 698 men from the Health Professionals Follow-up Study (1986-2008) and 83 644 women from the Nurses' Health Study (1980-2008) who were free of cardiovascular disease (CVD) and

1.23) and 1.21 (1.13-1.31) for CVD mortality and 1.10 (1.06-1.14) and 1.16 (1.09-1.23) for cancer mortality. We estimated that substitutions of 1 serving per day of other foods (including fish, poultry, nuts, legumes, low-fat dairy, and whole grains) for 1 serving per day of red meat were associated with a 7% to 19% lower mortality risk. We also estimated that 9.3% of deaths in men and 7.6% in women in these cohorts could be prevented at the end of fol-

jor lifestyle and dietary risk factors, the pooled hazard ratio (HR) (95% CI) of total mortality for a 1-serving-per-day increase was 1.13 (1.07-1.20) for unprocessed red meat and 1.20 (1.15-1.24) for processed red meat.

Departments of Nutrition (Drs Pan, Sun, Bernstein, Stampfer, Willett, and Hu) and Epidemiology (Drs Manson, Stampfer, Willett, and Hu), Harvard School of Public Health, and Channing Laboratory (Drs Sun, Stampfer, Willett, and Hu) and Division of Preventive Medicine (Dr Manson), Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; Wellness Institute of the Cleveland Clinic, Lyndhurst, Ohio (Dr Bernstein); and Department of Molecular Epidemiology, German Institute of Human Nutrition, Nuthetal, Germany (Dr Schulze)

of diabetes," cardiovascular disease (CC) and certain cancers. Several studies also suggest an elevated risk of mortality associated with red meat intake. However, most of these studies have been performed in populations with a particularly high proportion of vegatians (Such as Seventh- Day Adventists in

#### See Invited Commentary at end of article

the United States\* and several studies in Europe\*). A recent large cohort study\* will 10 years of follow-up found that a higher in take of total red meat and total processed mea was associated with an increased risk of mor tality. However, this study did not differentiate unprocessed from processed red mea and diet and other covariates were assessed at baseline only. Furthermore, to our knowledge of the contraction of the covariates were assessed at the covariates were assessed as the covariates were as the covariates wer

norts with repeated measures of det and up to 28 years of follow-up; the Health Professionals Follow-up Study (HPFS) and the Nurses' Health Study (NHS). We also estimated the associations of substituting other healthy protein sources for red meat with total and cause-specific mortality.

#### METHODS

#### STUDY POPULATIO

We analyzed data from 2 prospective cohort studies: the HPFS (initiated in 1986, n=51529 meraged 40-75 years) and the NHS (started in 1976 n=121700 women aged 30-55 years). Detailed descriptions of the cohorts are provided else where. <sup>7,8</sup> Questionnaires were administered biennially to collect and update medical, lifestyle and other health-related information, and the follow-up rates exceeded 90% in each 2-year cycle for both cohorts.

Annals of Internal Medicine, 12 maart 2012

ARCH INTERN MED PUBLISHED ONLINE MARCH 12, 2012 WWW.ARCHINTERNMED.CO

El

aded from www.archinternmed.com at Leiden University Medical Center, on March 14, 201

#### Research



Ratio measures in leading medical journals: structured review of accessibility of underlying absolute risks

Lisa M Schwartz, Steven Woloshin, Evan L Dvorin, H Gilbert Welch

Conclusion Absolute risks are often not easily accessible in articles reporting ratio measures and sometimes are missing altogether—this lack of accessibility can easily exaggerate readers' perceptions of benefit or harm.

cohort studies (62% v 21%; relative risk 3.0, 95% confidence interval 2.1 to 4.2) and for studies reporting crude compared with adjusted ratio measures (62% v 21%; relative risk 3.0, 2.1 to 4.3).

Conclusion Absolute risks are often not easily accessible in articles reporting ratio measures and sometimes are missing altogether—this lack of accessibility can easily exaggerate readers' perceptions of benefit or harm.

#### Introduction

For good reasons, ratio measures, such as relative risks and odds ratios, have become a common way to compare outcomes in two groups. For observational studies they serve as the central metric of the strength of association between exposure and outcome, a key criterion for establishing causality in classic epidemiology. For both observational and experimental studies they serve as a convenient mechanism to express the magnitude of an effect on baseline risk—a relative change. Furthermore, the ratio is often believed to be transportable—that is, the relative change may be applied to different populations with different baseline risks. Finally, they have the appealing feature of summarising two numbers (the risk in one group and the risk in the other) into one, which in turn facilitates comparisons of the effect of various exposures. But this feature of ratio measures also represents their major weakness, that the underlying absolute risks are concealed.

numbers when feasible.\* A more recent initiative organised in 2003, strengthening the reporting of observational studies in epidemiology (STROBE), does as well.\* We systematically determined how frequently the absolute risks comprising ratio measures are reported in the medical literature.

#### Methods

#### Search strategy

We searched Medline using the search terms rate ratio\*, relative ratio\*, relative risk\*, risk ratio\*, or odds ratio\* (the \* ensures that any suffix such as "s" is included), and identified 320 articles with ratio measures in the abstract published between 1 June 2003 and 1 May 2004 in six leading medical journals: Annals of Internal Medicine, BMJ, Journal of the American Medical Association, Journal of the National Cancer Institute, Lancet, and New England Journal of Medicine. We excluded 98 articles with study designs in which absolute risks might not be directly calculable (52 case-control studies, 46 meta-analyses). Thus the final sample consisted of 222 articles with study designs where absolute risks were directly calculable: 61 randomised trials, 161 cohort studies.

#### Article review process

We reviewed each article using a standardised data extraction form (see fig A on bmi.com). The coder began by recording the



The data extraction form and calculations carried out by coders are bmi.com

BMJ 16 december 2006

Table 2. All-Cause Mortality According to Red Meat Intake in the Health Professionals Follow-up Study and the Nurses' Health Study

	Frequency of Consumption Quintiles <sup>a</sup>			D.V.I	HR (95% CI) for a		
Variable	Q1	Q2	Q3	Q4	Q5	P Value for Trend	1-Serving-per-Day Increase
		Health	Professionals Follow	w-up Study			
Total red meat, servings per day <sup>b</sup>	0.25 (0.13-0.37)	0.61 (0.53-0.70)	0.95 (0.87-1.04)	1.36 (1.24-1.49)	2.07 (1.83-2.47)	NA	NA
Cases/person-years, No.	1713/151 212	1610/152 120	1679/151 558	1794/152 318	2130/151 315	NA	NA
Age-adjusted model	1 [Reference]	1.06 (0.99-1.14)	1.14 (1.06-1.21)	1.21 (1.14-1.30)	1.45 (1.36-1.54)	<.001	1.19 (1.16-1.23)
Multivariate model <sup>c</sup>	1 [Reference]	1.12 (1.05-1.20)	1.21 (1.13-1.30)	1.25 (1.16-1.34)	1.37 (1.27-1.47)	<.001	1.14 (1.10-1.17)
Unprocessed red meat, servings per day <sup>b</sup>	0.17 (0.07-0.24)	0.43 (0.37-0.47)	0.66 (0.58-0.73)	0.95 (0.87-1.04)	1.46 (1.29-1.67)	NA	NA
Cases/person-years, No.	1855/150 676	1722/149 097	1535/154 352	1819/150 925	1995/153 474	NA	NA
Age-adjusted model	1 [Reference]	1.06 (0.99-1.13)	1.00 (0.94-1.07)	1.15 (1.08-1.23)	1.34 (1.25-1.42)	<.001	1.22 (1.18-1.27)
Multivariate model <sup>c</sup>	1 [Reference]	1.11 (1.04-1.18)	1.14 (1.06-1.22)	1.20 (1.12-1.28)	1.29 (1.20-1.38)	<.001	1.17 (1.12-1.21)
Processed red meat, servings per day <sup>b</sup>	0.02 (0-0.07)	0.13 (0.10-0.14)	0.21 (0.20-0.26)	0.39 (0.34-0.46)	0.74 (0.64-1.00)	NA	NA
Cases/person-years, No.	1917/171 619	1395/131 069	1661/152 481	1717/152 128	2236/151 227	NA	NA
Age-adjusted model	1 [Reference]	0.99 (0.93-1.06)	1.13 (1.05-1.20)	1.14 (1.07-1.22)	1.38 (1.30-1.47)	<.001	1.34 (1.28-1.40)
Multivariate model <sup>c</sup>	1 [Reference]	1.06 (0.99-1.14)	1.15 (1.07-1.23)	1.18 (1.10-1.27)	1.27 (1.19-1.36)	<.001	1.18 (1.12-1.24)
			Nurses' Health Stud	dy			
Total red meat, servings per day <sup>b</sup>	0.51 (0.37-0.61)	0.85 (0.76-0.96)	1.14 (1.03-1.32)	1.49 (1.33-1.71)	2.17 (1.85-2.66)	NA	NA
Cases/person-years, No.	2946/438 326	2759/442 134	2658/439712	2872/440 329	3765/439 391	NA	NA
Age-adjusted model	1 [Reference]	1.07 (1.01-1.12)	1.09 (1.04-1.15)	1.24 (1.18-1.30)	1.61 (1.53-1.69)	<.001	1.30 (1.28-1.33)
Multivariate model <sup>c</sup>	1 [Reference]	1.08 (1.02-1.14)	1.11 (1.05-1.17)	1.18 (1.12-1.24)	1.24 (1.17-1.30)	<.001	1.11 (1.08-1.13)
Unprocessed red meat, servings per day <sup>b</sup>	0.37 (0.28-0.46)	0.61 (0.56-0.68)	0.86 (0.77-1.00)	1.13 (1.01-1.28)	1.64 (1.43-2.05)	NA	NA
Cases/person-years, No.	3079/441 041	2885/441 207	2545/439 306	2709/431 097	3782/447 240	NA	NA
Age-adjusted model	1 [Reference]	1.05 (1.00-1.11)	0.98 (0.93-1.03)	1.09 (1.03-1.14)	1.48 (1.41-1.55)	<.001	1.31 (1.28-1.35)
Multivariate model <sup>c</sup>	1 [Reference]	1.07 (1.01-1.12)	1.07 (1.01-1.12)	1.10 (1.05-1.16)	1.19 (1.13-1.25)	<.001	1.10 (1.06-1.13)
Processed red meat, servings, per day <sup>b</sup>	0.05 (0-0.11)	0.14 (0.13-0.16)	0.23 (0.21-0.28)	0.36 (0.33-0.42)	0.64 (0.56-0.87)	NA	NA
Cases/person-years, No.	3076/442 594	2799/420 403	2778/455 365	2814/441 369	3533/440 161	NA	NA
Age-adjusted model	1 [Reference]	1.06 (1.01-1.12)	1.10 (1.04-1.16)	1.18 (1.12-1.24)	1.49 (1.42-1.56)	<.001	1.61 (1.54-1.69)
Multivariate model <sup>c</sup>	1 [Reference]	1.04 (0.99-1.10)	1.08 (1.03-1.14)	1.14 (1.08-1.20)	1.20 (1.14-1.27)	<.001	1.21 (1.15-1.27)
			Pooled Results d				
Total red meat	1 [Reference]	1.10 (1.05-1.14)	1.15 (1.06-1.26)	1.21 (1.14-1.28)	1.30 (1.18-1.43)	<.001	1.12 (1.09-1.15)
Unprocessed red meat	1 [Reference]	1.08 (1.05-1.12)	1.10 (1.03-1.17)	1.15 (1.05-1.25)	1.23 (1.14-1.34)	<.001	1.13 (1.07-1.20)
Processed red meat	1 [Reference]	1.05 (1.00-1.09)	1.11 (1.04-1.18)	1.15 (1.11-1.20)	1.23 (1.16-1.30)	<.001	1.20 (1.15-1.24)

Sterfte- risico	Sterfte per 1000 mannen
I	231
I,I2	258
1,21	279
1,25	288
1,37	316
	risico  I,12 I,21 I,25

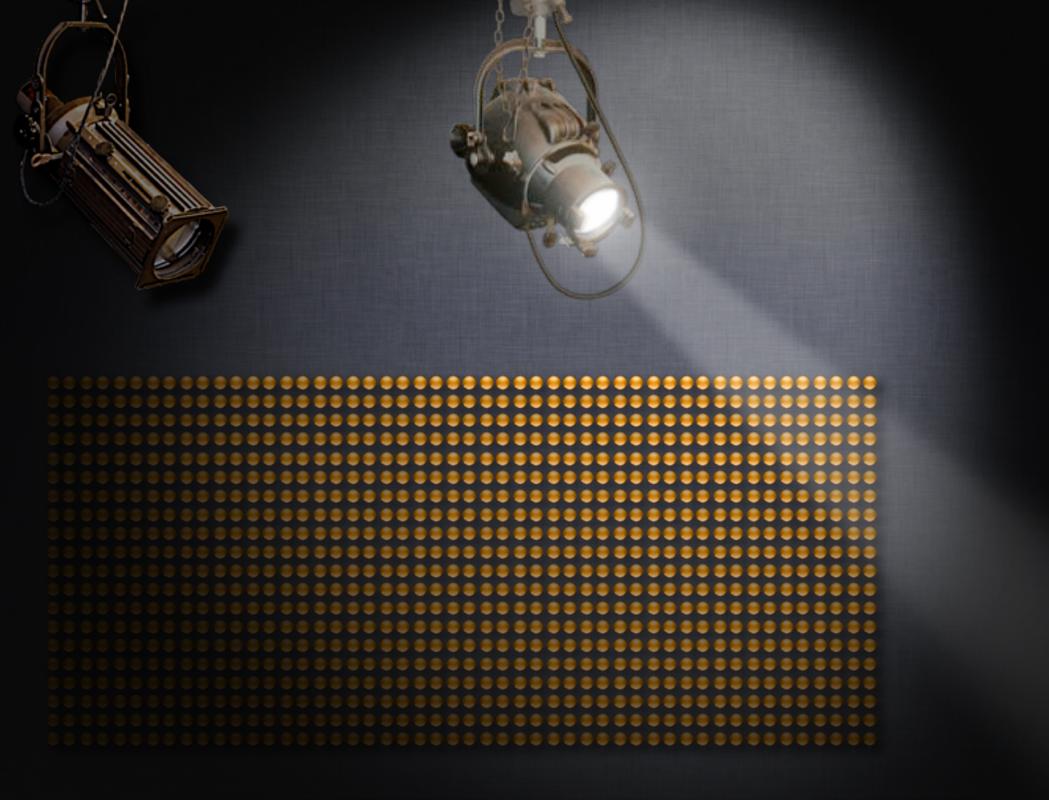
Sterfte onder mannen na 22 jaar

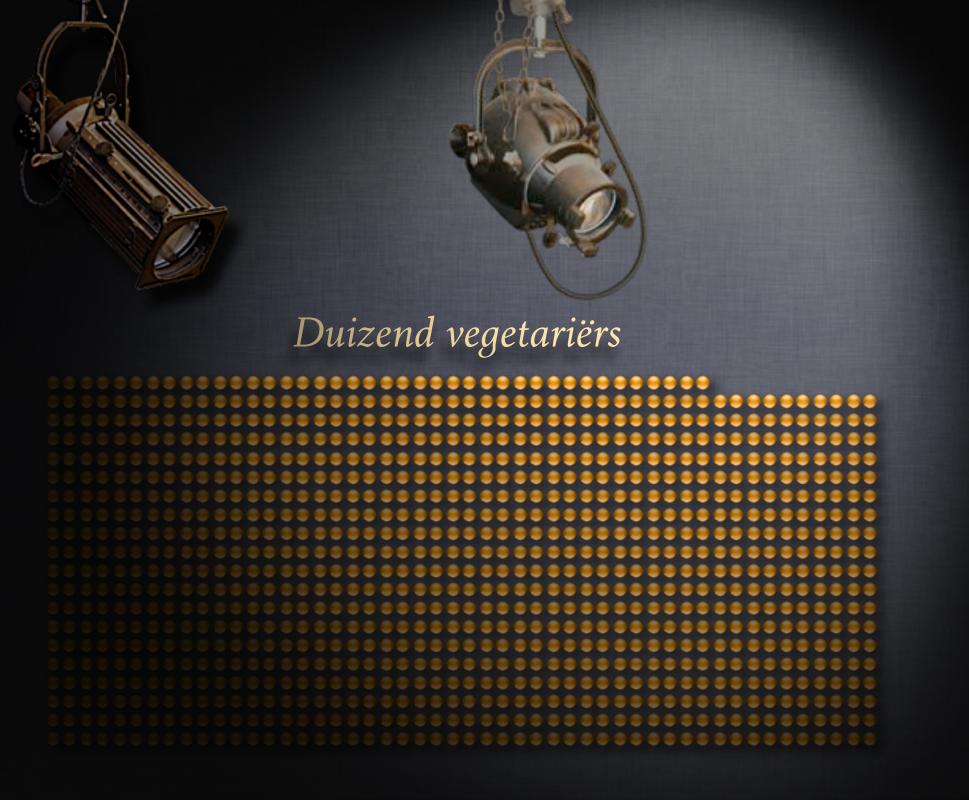


Met hoeveel procent verhoogt een roker zijn kans op longkanker?

Porties per dag	Sterfte- risico	Sterfte per 1000 per jaar
6,3	I	II
0,6	I,I2	12
I,O	I,2I	13
I,4	I,25	13
<b>2,</b> I	1,37	14

Sterfte onder mannen na 1 jaar







#### Cardioxascular mortality in Dotch men during 1996 European foodsall championship: longitudinal population study

Basel E West, Michael J. Born, New Y. Torn, Darlands T. Godding,

Militari I I--Service - I Store In Sec. Statement & Spinner.

Physiological Science Space of Service and Emperorant Saddle sand because term is set or own that I Viggers are to recognise indicates and stoke. Bridge I coughwheel mady of secondity properly If Jane 1996 the die the Data's bested was vondispersion to harpen to def-Company of the calls on 17 has no company with the liter than before and plice the matrix and inthe rates partially a rived quart reed. Belling Technicals

Make the hope to be upon the common result. See 1996.

Main metrode desperant follows contain and provides for a consum from those and make Emple World's from corners bear down and make on borough's open as he for it's such behalve the 17th William Material States in the St. \$1000. No view view to expenditive way observed the STATE OF THE PARTY AND ADDRESS OF THE PARTY ADDRESS OF THE PARTY AND ADDRESS OF THE PARTY ADDRESS OF THE PARTY AND ADDRESS OF THE PARTY AND ADDRESS OF THE PARTY AND ADDRES store carbonards itself-estand in the land for much

Employed Experient specing electrons protein a safficient broken more in edges magnessade radictorate form. To different between sengal weapen explore for the investigation.

#### Introduction:

as approached influences and stroke line proched Secretary attention to recome proc. There played at body street are by a conditional data across fact; abided on, and recently have been proposed at positive regarding factors for exposential defection. and make." These factors probably not us for find

For colo of visual factors in sever reactife more such.

toy it a separate of cours including the gradual Arrefrances of abstractional leasest

A few separa have suggested that classes that before seen on large repulser of people in federal area, such a composite and the long become the incidence of last and non-last exponented inflations: at the propolation best." The impedicated that a big. area shall continue around of the proposed bigget. factors at one point in time for a longe propriation. could have a place or market of course to condition. describe become in natural embermoda para cales. Novalido, in a company like the Northerlands, MC improved builted march of the national trans could case a continuous of incomed month and providence from the foreign and providing process. ing and concerns exacting on a population kind."

Service F. June and H. Jane 1986 the European hotel despends to part a highest file. Noticebook selement to the general hosts and placed. Property of 19 has the spark production and all times com also man been see? I saw took on penalty light-According to the colony positioned by the Dhanks broad aring broadense," about 95 celling garger wanted for man's tip represent your lifts of the Drawburgstate and Philosophica 1996.

#### Mindreds

We produced controller data from the Death controll. become for materials for flow 1995, data 1995, and flow (All The day comprised up and we specific resident of fronts, and study from a set operated and contract. and make \$100 Donates \$100 to and \$1000.

No compared the anades of third cital store team tally and first has reproduced interest or bridge on its day of the booked special CD James 1990; with for cares market of deaths on the patrocking and MA lowing the days he man and somes repeated Withthe city with 10% confidence beautiful team. street in addition to compared the data with how how corresponding periods in 1985 and 1985.

#### Ersales

To have done to made of holls from all course and commanded information on market thinking (Coll.) Street. THE R CO. LANSING MICHIGAN PROPERTY AND ADDRESS OF THE PARTY AND ADDRES name to be comed to make on 15 June 1778 a 1780. NAME AND POST OFFICE ADDRESS OF THE PARTY. I Fill and have be worse (CM) a Still Course Wilder Add Add A Thin & Mr. brown, morable from some at AN INVESTIGATION OF THE PARTY NAMED IN COLUMN where out 130, 100 to 100 to the day of the last.



BMJ25-31 december 2000



Extract C. C. Stiffing 1 No. Spring sect., Blood 11

store carbonoute thate-econoline to body of

for much



**Results** Mortality from coronary heart disease and stroke was increased in men on the day of the match (relative risk 1.51, 95% confidence interval 1.08 to

Program on 19 June. The search constraints and all filling.

count after more from any Transa was no person for the

According to the colong particles for the Death

broad uning Departmen," about 757 celling party wanted do mark the represent one diffe of the



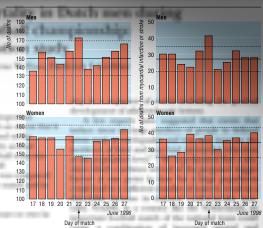
BMJ25-31 december 2000 (21 cases) on the day after the match, but not below the lower bound of the 95% confidence interval.

In women, no clear difference in numbers of deaths from myocardial infarction and stroke was observed (38 v 34.1 cases; relative risk 1.11,0.8 to 1.56). Analyses of the same periods in 1995 and 1997 showed no significant increases in cardiovascular mortality on 22 June compared with the respective previous and following five days.

#### Discussion

Mortality from acute myocardial infarction and stroke was increased in Dutch men on the day of an important football match compared with the preceding and following days. Compared with the average mortality in the preceding and following days, about 14 additional fatal events occurred; this is an increase of around 50%. Our findings indicate that watching an important football match may be stressful. It further supports the view that the mental and other stressors of such an event may trigger acute myocardial infarction or stroke.

No corresponding increase in mortality occurred in women. This could mean that fewer women were exposed to the triggers (for example, because of less interest in football or lower alcohol consumption) or that they are less vulnerable to their effect. Less exposure seems the more likely explanation, although Tofler et al found that men were more likely to report triggers before myocardial infarction than women.\(^1\) The exact mechanism behind this difference is still unclear.



No corresponding increase in mortality occurred in women. This could mean that fewer women were exposed to the triggers (for example, because of less

role of emotional or mental stress, physical activity, and heavy alcohol ingestion has been established mainly by retrospective questioning of people who have had non-fatal cardiovascular events. The finding of ingregory



**Results** Mortality from coronary heart disease and stroke was increased in men on the day of the match (relative risk 1.51, 95% confidence interval 1.08 to



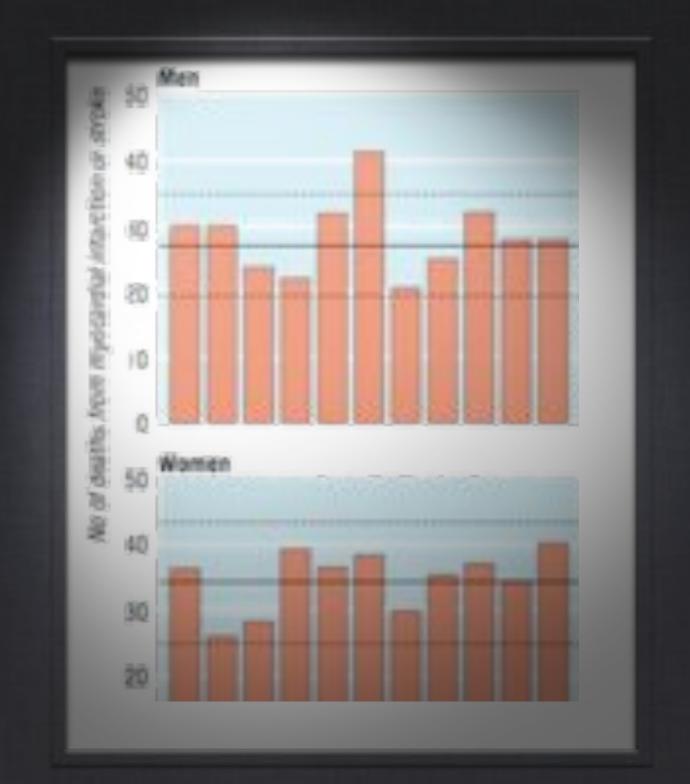
nduce stress on a large number of people in a defined earthquakes and war, have been reported to increase the fatal and non-fatal myocardial infarction

#### study add:

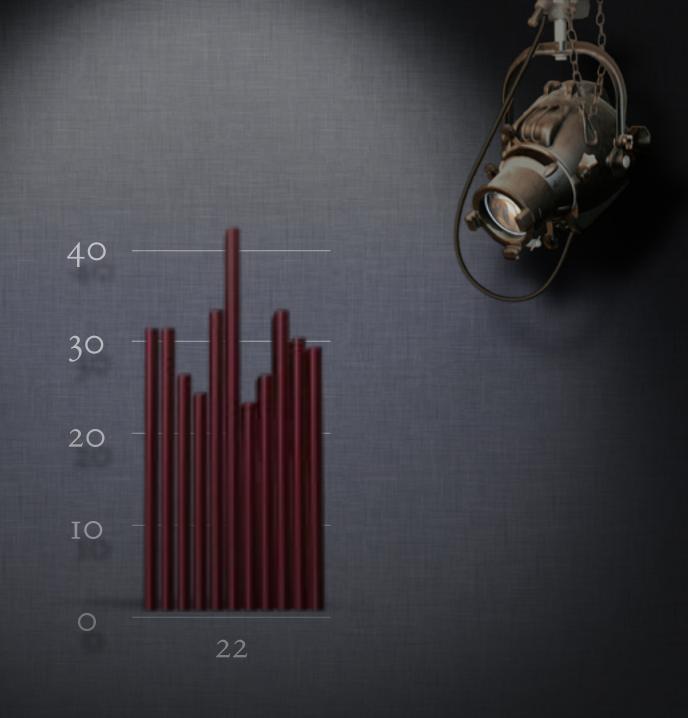
lity from myocardial infarction and stroke was increased on n important football match of the Dutch national team, cause of increased stress

women was unaffected

The effect of a single match is detectable in national mortality data



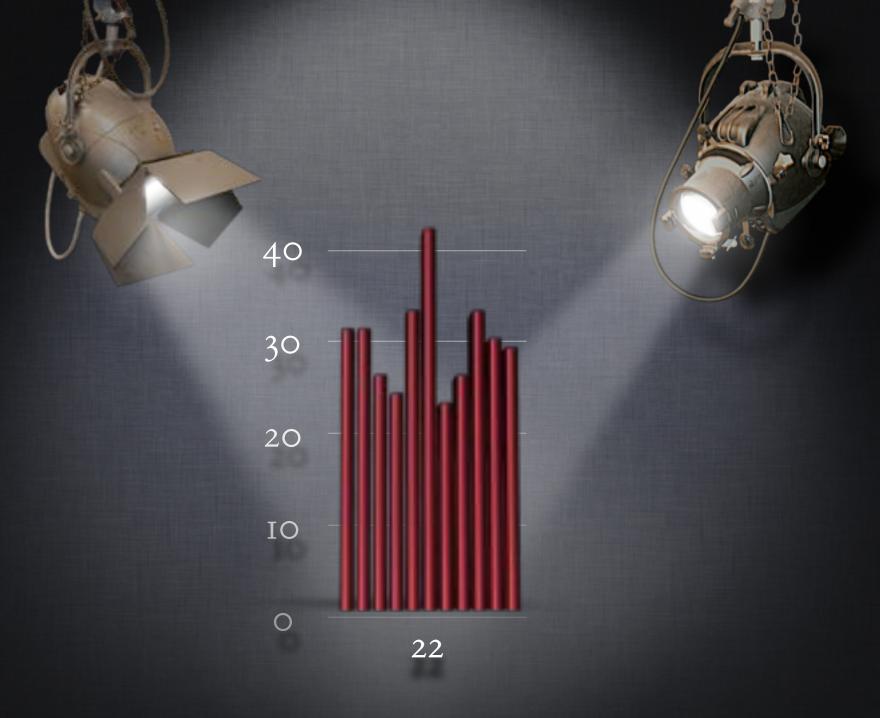




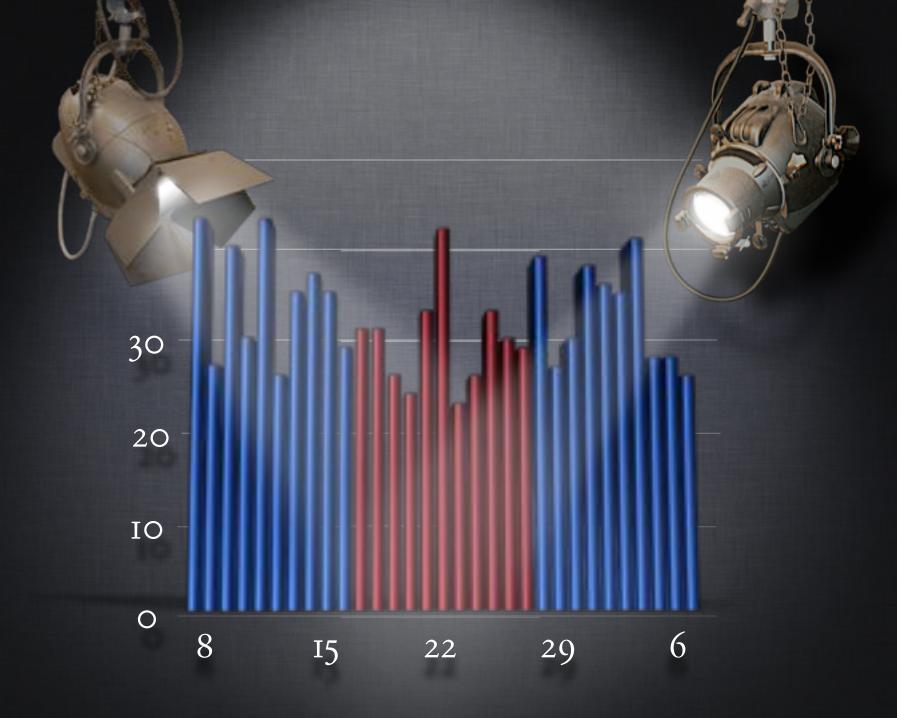
Hartdoden Nederlandse middelbare mannen, juni 1996

# Deburnalistieke toets

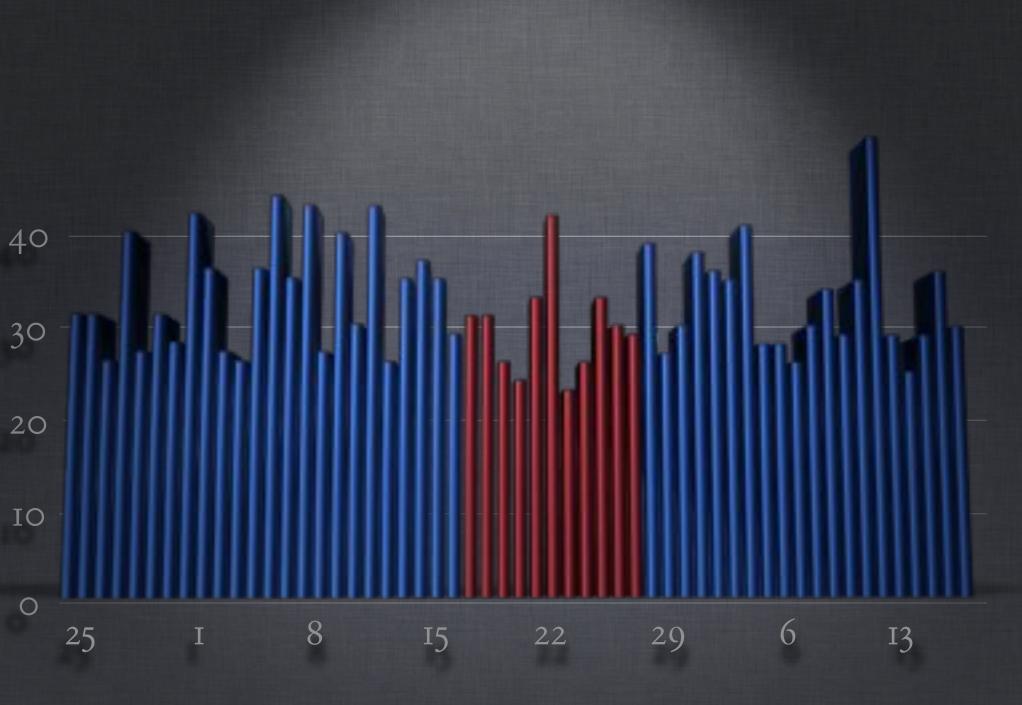
Bedenk ten minste drie bezwaren tegen deze conclusie



Hartdoden Nederlandse middelbare mannen, juni 1996



Hartdoden Nederlandse middelbare mannen, zomer 1996



Hartdoden Nederlandse middelbare mannen, zomer 1996

# De journalistieke toe

- Kan het kloppen?
- Waarom deze cijfers, en alleen deze cijfers?
- Hoe hebben ze dat gemeten — en zo precies gemeten?

# De journalistieke toe

- Hoe moet ik me dit voorstellen?
- Zijn er geen andere verklaringen te bedenken?
- Wat zegt het gezond verstand?