

When One Depends on the Other

Reporting of Interaction in Case-Control and Cohort Studies

Mirjam J. Knol,^{a,b,c} Matthias Egger,^{c,d} Pippa Scott,^c Mirjam I. Geerlings,^a and Jan P. Vandenbroucke^c

Background: Interaction refers to the situation in which the effect of 1 exposure on an outcome differs across strata of another exposure. We did a survey of epidemiologic studies published in leading journals to examine how interaction is assessed and reported.

Methods: We selected 150 case-control and 75 cohort studies published between May 2001 and May 2007 in leading general medicine, epidemiology, and clinical specialist journals. Two reviewers independently extracted data on study characteristics.

Results: Of the 225 studies, 138 (61%) addressed interaction. Among these, 25 (18%) presented no data or only a *P* value or a statement of statistical significance; 40 (29%) presented stratum-specific effect estimates but no meaningful comparison of these estimates; and 58 (42%) presented stratum-specific estimates and appropriate tests for interaction. Fifteen articles (11%) presented the individual effects of both exposures and also their joint effect or a product term, providing sufficient information to interpret interaction on an additive and multiplicative scale. Reporting was poorest in articles published in clinical specialist articles and most adequate in articles published in general medicine journals, with epidemiology journals in an intermediate position.

Conclusions: A majority of articles reporting cohort and case-control studies address possible interactions between exposures. However, in about half of these, the information provided was unsatisfactory, and only 1 in 10 studies reported data that allowed

readers to interpret interaction effects on an additive and multiplicative scale.

(*Epidemiology* 2009;20: 161–166)

Interaction refers to the situation in which the effect of 1 exposure on an outcome differs across strata of another exposure. Other terms that have been used to describe interaction include “effect modification,” “effect measure modification,” or “synergism” and “antagonism.”

From a statistical point of view, interaction refers to the necessity of a product term in a statistical model.¹ Depending on the form of the statistical model, interaction is examined on an additive scale (for example in linear regression) or on a multiplicative scale (for example in logistic regression). In epidemiology, there has been a long-standing debate on whether the scale should be determined by the statistical model that fits best or whether interaction should be assessed on an additive scale irrespective of the underlying statistical model.^{2–8} It has been argued that the additive scale is more appropriate to assess “biologic interaction,” which is implied by terms such as synergism and antagonism.^{6,7}

An important argument for using the additive scale is that it fits with the sufficient-component concept of causality.^{7,8} The presence or absence of interaction on an additive scale does not, however, indicate a particular disease mechanism. The sufficient-component theory of causation is a deliberate abstraction, independent of an underlying disease mechanism.^{7,9,10} This may be one reason why not all researchers subscribe to the notion that interaction should always be presented on an additive scale. It has been recommended that reports on interaction should provide sufficient information so that readers can interpret interaction on an additive and on a multiplicative scale.^{2,11} Such reporting would allow readers to interpret interaction additively (ie, in a sufficient component-cause framework) whatever the opinion of the authors. This can be done by presenting the individual effects of both exposures and their joint effect, each relative to the group not exposed to either risk factor.² It can also be done by presenting the full statistical model: the individual effects of both exposures and their product term, which allows readers to recalculate the joint effect.^{7,12,13} Although much has been written about interaction in methodologic papers, little is

Submitted 6 February 2008; accepted: 22 April 2008; posted 21 November 2008.

From the ^aJulius Center for Health Sciences and Primary Care, University Medical Center Utrecht; ^bDepartment of Pharmaco-epidemiology and Pharmacotherapy, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Utrecht, Netherlands; ^cInstitute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland; ^dDepartment of Social Medicine, University of Bristol, UK; and ^eDepartment of Clinical Epidemiology, Leiden University Medical Center, Leiden, Netherlands.

Supported by a grant from the Prince Bernhard Cultural Foundation, an unrestricted grant from Novo Nordisk and the Scientific Institute of Dutch Pharmacists (WINAp) and a VIDI grant from the Netherlands Organization for Scientific Research (NWO: project 917-66-311).

J.P. Vandenbroucke currently is an Academy Professor of the Royal Netherlands Academy of Arts and Sciences.

 Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

Editors' Note: An editorial related to this article appears on page 159.

Correspondence: M. J. Knol, University Medical Center Utrecht, Julius Center for Health Sciences and Primary Care, PO Box 85500, 3508 GA, Utrecht, Netherlands. E-mail: m.j.knol@umcutrecht.nl.

Copyright © 2008 by Lippincott Williams & Wilkins

ISSN: 1044-3983/09/2002-0161

DOI: 10.1097/EDE.0b013e31818f6651

known about epidemiologic practice. How frequently is interaction examined in epidemiological studies? What approach do authors take when investigating and reporting interaction between exposures? We conducted a survey of case-control and cohort studies published in leading journals to examine how interaction is currently assessed and reported in cohort and case-control studies.

METHODS

Selection of Articles

We examined case-control and cohort studies published in 5 leading general medicine journals (*Annals of Internal Medicine*, *British Medical Journal*, *Journal of the American Medical Association*, *Lancet*, *New England Journal of Medicine*), 5 leading epidemiology journals (*American Journal of Epidemiology*, *EPIDEMIOLOGY*, *International Journal of Epidemiology*, *Journal of Clinical Epidemiology*, and *Journal of Epidemiology and Community Health*), and 10 leading clinical specialist journals (*American Journal of Respiratory and Critical Care Medicine*, *Archives of General Psychiatry*, *Arthritis and Rheumatism*, *Blood*, *Circulation*, *Clinical Infectious Diseases*, *Diabetes Care*, *Journal of American Geriatrics Society*, *Journal of the National Cancer Institute*, and *Pediatrics*). We selected studies in a literature search performed at the end of March 2007 combining the journal names with the MESH term “case-control studies” and the MESH term “cohort studies.” Subsequently, we identified eligible studies starting with the issues published in March 2007 and went backward in time until we identified 150 eligible case-control studies and 75 eligible cohort studies: 10 case-control studies from each general medicine and each epidemiology journal, and 5 case-control studies from each clinical specialist journal; 5 cohort studies from each general medicine and each epidemiology journal, and 2 or 3 cohort studies from each clinical specialist journal. Articles that were published electronically ahead of print were included.

We selected more case-control studies than cohort studies because we included the case-control studies also in a separate study on the interpretation of the odds ratio in case-control research.¹⁴ Case-crossover studies and studies that did not report any measure of association were excluded. Finally, we considered only original articles and short reports.

Data Extraction

We developed a standardized data extraction form. The extraction form was piloted on 6 articles not included in the study and subsequently revised. We extracted general items including the number of study participants, main exposure and condition studied, and 4 interaction-specific items. First, we assessed whether interaction was addressed. Second, we extracted what terms the authors used to describe interaction. Third, we assessed how the interaction was presented, for example in stratified analyses or by reporting a *P* value. We

TABLE 1. Two Types of Stratification: (1) the Presentation of Effect Estimates of Exposure and Outcome in Strata of the Suspected Effect-Modifying Exposure; (2) the Presentation of Individual Effects of Both Exposures and Their Joint Effect, Each Relative to No Exposure

Exposure	Example 1 RR (95% CI)	Example 2 RR (95% CI)
No smoking		
No obesity	1.0 (reference)	1.0 (reference)
Obesity	1.8 (1.4–2.4)	1.8 (1.4–2.4)
Smoking		
No obesity	1.0 (reference)	1.3 (1.0–1.8)
Obesity	1.7 (1.2–2.3)	2.2 (1.6–3.0)

In this hypothetical example, obesity is the risk factor of interest, and smoking, the suspected effect modifier.

distinguished 2 types of stratification: (1) the presentation of effect estimates of exposure and outcome in strata of the suspected effect-modifying exposure; (2) the presentation of individual effects of both exposures and their joint effect, each relative to no exposure. Table 1 illustrates the 2 types of stratification with a hypothetical example. Fourth, we assessed what statistical tests for interaction were reported, for example a Wald test or likelihood ratio test. The Wald test assesses whether the regression coefficient of the product term is statistically different from zero. The likelihood ratio test assesses whether the model with the product term provides a better fit than the model without the product term.

Two reviewers (M.J.K. and P.S.) independently assessed all articles. Discrepancies were discussed by the 2 reviewers; if necessary, a third person (J.P.V. or M.E.) was consulted to reach consensus.

Levels of Reporting Interaction

We defined 4 levels of reporting interaction, where each level gave more information about the degree and direction of the interaction. Level 1 consists of only a *P* value or statement of statistical significance, or only a quantitative description that there was no interaction. In level 2, separate effect estimates and confidence intervals are provided across strata of the suspected effect modifying exposure but without meaningful comparison of stratum-specific effect estimates. In level 3, there are effect estimates and confidence intervals across strata of the other factor and a *P* value or statement of statistical significance based on an appropriate test for interaction. Finally, in level 4, there is sufficient information to interpret interaction on an additive and on a multiplicative scale. This could be done either by the presentation of the individual effect estimates and the joint effect estimate using 1 reference category or by the presentation of the full model, that is, the individual effect estimates and the effect estimate of the product term. Additional measures of interaction may be presented, including the synergy index of multiplicativity

TABLE 2. Study Characteristics of 225 Case-Control and Cohort Studies by Type of Journal

	General Medicine Articles (n = 75) No. (%)	General Epidemiology Articles (n = 75) No. (%)	Clinical Specialist Articles (n = 75) No. (%)
Country of authors ^a			
United States	42 (56)	32 (43)	44 (59)
Europe except UK	27 (36)	30 (40)	26 (35)
United Kingdom	17 (23)	13 (17)	13 (17)
Other	22 (29)	18 (24)	17 (23)
Country of study participants ^a			
United States	33 (44)	23 (31)	32 (43)
Europe except UK	20 (27)	30 (40)	25 (33)
United Kingdom	15 (20)	13 (17)	10 (13)
Other	20 (27)	13 (17)	16 (21)
No. subjects ^b	3570 (57–383,862)	3078 (100–10,766,971)	1206 (33–1,380,000)
Exposure category ^a			
Treatment	29 (39)	9 (12)	15 (20)
Prevalent medical conditions	9 (12)	9 (12)	19 (25)
Lifestyle factors	9 (12)	15 (20)	10 (13)
Other	32 (40)	46 (57)	39 (51)
Outcome category ^a			
Cardiovascular disease	26 (35)	12 (16)	20 (27)
Cancer	6 (8)	23 (31)	14 (19)
All-cause mortality	12 (16)	11 (15)	6 (8)
Infectious disease	12 (16)	4 (5)	13 (17)
Other	20 (27)	26 (35)	23 (31)

^aMore than 1 possible in 1 article.^bMedian (range) presented.

or additivity, or the relative excess risk due to interaction.^{6,12} Published examples of the various levels of reporting are shown in eTable 1 (<http://links.lww.com/A716>).

Data Analysis

Frequencies and summary statistics were calculated stratified by journal type (general medicine, epidemiology, and clinical specialist journals) and study type (case-control or cohort studies).

RESULTS

Our literature search for case-control studies produced 4647, 3351, and 6508 hits in the general medicine, epidemiology, and clinical specialist journals, respectively. Based on this search, we identified the 50 most recent eligible case-control studies for each journal type. The literature search for cohort studies resulted in 13,856, 4986, and 15,348 hits, respectively. Again, we selected the 25 most recent eligible cohort studies for each journal type. The publication date of the selected articles ranged from May 2001 to April 2007 in the general medicine journals (median: June 2006), from October 2002 to May 2007 in the general epidemiology journals (median: October 2006), and from August 2004 to April 2007 in the clinical specialist journals (median: January 2007). Fourteen of the 225 articles (6%) were short reports; 6 in general medicine journals,

4 in general epidemiology journals, and 4 in clinical specialist journals. The citations for these 225 articles are presented in eTable 2 (<http://links.lww.com/A715>).

Table 2 presents the characteristics of the 225 studies, stratified by journal type. Most authors and study participants came from the United States or Europe. The largest studies were published in general medicine journals, closely followed by studies published in epidemiology journals. Studies published in clinical specialist journals were smaller. As expected, cohort studies included more individuals than case-control studies (data not shown). The most frequent exposures were treatments, followed by prevalent medical conditions and lifestyle factors. Cardiovascular disease, cancer, and all-cause mortality were the most frequently studied outcomes.

Table 3 describes details on the reporting of interaction. About two third of the studies (138 articles, 61%) examined interaction between exposures. In 12 of these studies (9%), mainly published in epidemiology journals, this was stated as 1 of the objectives of the study. An additional 5 studies (4%), of which 4 were published in general medicine journals, explicitly stated that the interaction analyses were prespecified. In all journals, the most frequently used terms were “interaction” and “effect (measure) modification” (101 of 138 articles, 73%). The term “subgroup analysis” was used in 18

TABLE 3. Reporting on Interaction in 225 Case-Control and Cohort Studies by Type of Journal

	General Medicine Articles (n = 75) No. (%)	General Epidemiology Articles (n = 75) No. (%)	Clinical Specialist Articles (n = 75) No. (%)
Interaction examined	46 (61)	50 (67)	42 (56)
Terms used ^a			
Interaction	29 (63)	19 (38)	24 (57)
Effect (measure) modification	10 (22)	12 (24)	7 (17)
Subgroup analysis	9 (20)	5 (1)	4 (10)
Stratification	3 (7)	4 (8)	4 (10)
Synergy	2 (4)	3 (6)	1 (2)
No specific term used	3 (7)	9 (18)	5 (12)
Presentation of interaction ^a			
Stratum-specific effect estimates ^b	33 (72)	36 (72)	29 (69)
Individual and joint effect estimates ^b	4 (9)	6 (12)	4 (10)
P	22 (48)	13 (26)	16 (38)
Statement on statistical significance	8 (17)	13 (26)	10 (24)
Product term	2 (4)	0 (0)	1 (2)
Synergy index	2 (4)	0 (0)	1 (2)
No data presented	4 (9)	3 (6)	1 (2)
Statistical tests used			
Wald test for product term	11 (24)	11 (22)	6 (14)
Likelihood ratio test	6 (13)	8 (16)	9 (21)
Heterogeneity test	1 (2)	1 (2)	0 (0)
Unclear	14 (30)	7 (14)	13 (31)

^aMore than 1 possible in 1 article, percentages relate to studies that examined interaction.^bSee Table 1 for example.**TABLE 4.** Different Levels of Reporting on Interaction in 225 Case-Control and Cohort Studies by Type of Journal

Level of Reporting ^a	Total (n = 138) No. (%)	General Medicine Articles (n = 46) No. (%)	General Epidemiology Articles (n = 50) No. (%)	Clinical Specialist Articles (n = 42) No. (%)
Level 1	25 (18)	8 (17)	8 (16)	9 (21)
Level 2	40 (29)	10 (22)	16 (32)	14 (33)
Level 3	58 (42)	23 (50)	20 (40)	15 (36)
Level 4 ^b	15 (11)	5 (11)	6 (12)	4 (10)

^aSee eTable 1 (<http://links.lww.com/A716>) and main text for definitions and examples of levels of reporting.^bThree of studies of level 4 mentioned the additive scale when discussing the examined interaction.

studies (13%), and “stratification” or “stratified analysis” in 11 studies (8%). “Synergy,” “combined effect,” or “joint effect” was mentioned in 6 studies (4%). Seventeen studies (12%) did not use an explicit term.

The most frequent approach to reporting interactions (98 of 138 articles, 71%) involved the presentation of stratum-specific effect estimates (Table 1, example 1). Only 14 studies (10%) reported individual effects of both exposures and their joint effect (Table 1 example 2), and few studies reported product terms or a synergy index. No study presented the relative excess risk due to interaction or the attributable proportion due to interaction. *P* values or statements on statistical significance were reported in 82 (59%)

articles, but the statistical test was often unclear. One study incorrectly examined interaction by assessing overlap of the confidence intervals around the effect estimates in the strata. Another possible invalid interpretation is to conclude that interaction is present because the *P* value in 1 stratum is statistically significant but the *P* value in the other stratum is not. None of the studies explicitly reported this method, although 3 of the studies in which the statistical test was unclear may have used this approach.

When analyzing the data across “levels of reporting interaction” defined in the methods section, 25 articles (18%) were categorized as level 1 and 40 articles (29%) met criteria for level 2 (Table 4). Fifty-eight articles (42%) presented

effect estimates in strata of the suspected effect modifier and provided a *P* value or statement of statistical significance (level 3). Only a few articles (15, 11%) presented individual effect estimates of both exposures and their joint effect, or a product term from a statistical model (level 4). Three of these studies mentioned the additive scale when discussing interaction. The first study presented the synergy index for additive interaction in their results.¹⁵ A second study mentioned in the discussion that the observed interaction was more than could be expected under an additive or multiplicative model.¹⁶ The third study stated (also in the discussion) that some of the interactions were consistent with an additive model and some were consistent with a multiplicative model.¹⁷ Two other studies, which were both case-control studies, presented a synergy index of multiplicativity.^{18,19} Of note, among the studies in level 4, 6 did not provide a confidence interval around the synergy index or a *P* value for interaction.

Few differences were evident between groups of journals. Interestingly, the term interaction was used somewhat less frequently in epidemiology journals than in general medicine and clinical specialist journals (Table 3). Papers in epidemiology journals provided more complete descriptions of the statistical test used to assess interaction. Poorer reporting (levels 1 and 2) was most prevalent in articles published in clinical specialist articles (23 of 42 articles, 54%), whereas better reporting (levels 3 and 4) was most common in articles published in general medicine journals (28 of 46 articles, 61%). Studies published in epidemiology journals were intermediate (Table 4).

DISCUSSION

This survey of current practice regarding how interaction is assessed and reported in epidemiologic studies found that interaction between exposures is addressed in a majority of studies published in leading journals. However, in about half of the studies, the approach used or information reported was inadequate. In particular, brief descriptive statements or the presentation of stratum-specific estimates with no meaningful comparison across the strata were common. Only 1 in 10 studies reported individual effects of the 2 exposures and their joint effect or product term, thus providing sufficient information to interpret interaction on an additive as well as a multiplicative scale. Reporting on interaction was not more comprehensive in epidemiology journals compared with general medicine or clinical specialist journals.

Our study was based on articles published in leading journals and our results may therefore not be applicable to cohort or case-control studies published in less prominent journals. We did not assess whether the reported interaction would have been important from a clinical or public health perspective, and we did not assess whether the authors distinguished important from unimportant interactions. Many authors may feel that if little evidence for an interaction was

found, it is unnecessary to report greater detail. Furthermore, material may have been removed due to space constraints, word limits, or editorial intervention. A study of reports of clinical trials submitted to the *British Medical Journal* and later published either in the *British Medical Journal* or another journal does not, however, support this hypothesis: the number of tables and figures did not change markedly between submission and publication.²⁰

Pocock et al²¹ recently reviewed the analysis and reporting of epidemiologic research published in epidemiology journals or general and specialist medical journals. In line with our findings, a large proportion of articles (43 of 73, 59%) included subgroup analyses; the majority of these claimed differences across groups. Two surveys on reporting of clinical trials found that authors frequently made the invalid interpretation that interaction is present when effect estimates reach conventional levels of statistical significance (ie, $P < 0.05$) in 1 subgroup but not another (9 of 17 studies²²; 13 of 35 studies²³). Surprisingly, in our survey, only 3 studies may have used this method. Apparently, researchers in observational epidemiology are more aware of the shortcomings of this method of than researchers in the field of clinical trials.

In this survey, 17 of the 138 articles that reported interaction (12%) explicitly stated that the examination of interaction was an objective of the study or that the interaction analyses were prespecified. Some studies mentioned the assessment of interaction in the introduction, others in the methods, or only in the results. Analyses that were not the original aim of the study but arose during data analysis (eg, because of additional subject matter knowledge or new literature) may often be useful and provide important insights. However, it has been advised that such analyses should be described accordingly.^{11,23} The description may give information whether the original data collection (assessment of exposure and confounders relevant to the exposure) was adequate for a particular new hypothesis, and it allows the reader to assess the implications of multiple comparisons in the analysis.¹¹ To some readers, this is important information because they believe that data on a hypothesis that was specified beforehand is more credible, a view commonly held for randomized trials.²⁴ Readers with other views about inference, mainly in etiologic research, will pay less attention to whether the hypothesis was prespecified because they will argue that the prior probability of the hypothesis (whether known before or after looking at the data) is more important, and that replications will tease out which findings stand the test of time.^{25,26}

The reporting of interaction on an additive scale is very uncommon. Only 3 of the 138 studies mentioned the additive scale for the interpretation of interaction, and one of these calculated a synergy index for additivity. No study used the relative excess risk due to interaction or the attrib-

utable proportion due to interaction. In general, the presence of similar odds ratios across strata of a strong risk factor suggests strong interaction on an additive scale. We noticed that authors often do not consider this possibility and its potential public health impact. Interaction on the additive scale may reflect more appropriately the causal structure of interactions.^{3–5,7,8,11} Also, the additive scale, which uses absolute risks, may also be more appropriate for public health and clinical decision making.^{11,27} These concerns do not seem to have affected the practice of applied epidemiology.

One little-known drawback of calculating indices for additive interaction from multiplicative models is that for case-control studies, this calculation is exact only for tables or models that do not include other variables (eg, confounders) over which the index might vary.^{28,29} If the relative excess risk due to interaction is different among, say, men and women, the calculation from a logistic model that includes sex as an additional covariate will yield only an approximation of the sex-specific relative excess risks due to interaction. The synergy index for additivity seems to be more resistant to this problem: a simulation study showed it was stable in most situations and more often stable than the relative excess risk due to interaction.²⁹ In cohort studies, the problem does not arise if additive risk models are fitted,²⁸ which was not done in any paper in our survey.

Our study demonstrates that the reporting in epidemiological studies on interaction is often inadequate for a full assessment of additive or multiplicative interaction. Recommendations put forward by the STROBE initiative^{11,30} address several aspects relevant to interaction and subgroup analyses. STROBE asks authors to describe any methods used to examine subgroups and interactions and to report which analyses were planned in advance to allow readers to judge the implications of multiplicity. Botto and Khoury² proposed that authors should present individual effects of exposures and joint effects, each relative to no exposure, as the most informative approach to presenting the data. A less intuitive alternative is the presentation of the full model, including the interaction term. Only 1 in 10 articles (15 of 138) in our survey had either of these 2 presentations. Greater awareness among epidemiologists of the utility of more complete presentation of interaction results, such as advocated by publication guideline like STROBE,³⁰ may encourage a change in this practice.

REFERENCES

- Ahlbom A, Alfredsson L. Interaction: a word with two meanings creates confusion. *Eur J Epidemiol*. 2005;20:563–564.
- Botto LD, Khoury MJ. Commentary: facing the challenge of gene-environment interaction: the two-by-four table and beyond. *Am J Epidemiol*. 2001;153:1016–1020.
- Greenland S, Poole C. Invariants and noninvariants in the concept of interdependent effects. *Scand J Work Environ Health*. 1988;14:125–129.
- Rothman KJ. Synergy and antagonism in cause-effect relationships. *Am J Epidemiol*. 1974;99:385–388.
- Rothman KJ, Greenland S, Walker AM. Concepts of interaction. *Am J Epidemiol*. 1980;112:467–470.
- Rothman KJ. Interactions between causes. In: *Modern Epidemiology*. Boston, Toronto: Little Brown and Company; 1986:311–326.
- Greenland S, Rothman KJ. Concepts of interaction. In: *Modern Epidemiology*. 2nd ed. Philadelphia: Lippincott-Raven Publishers; 1998:329–342.
- VanderWeele TJ, Robins JM. The identification of synergism in the sufficient-component-cause framework. *Epidemiology*. 2007;18:329–339.
- Siemiatycki J, Thomas DC. Biological models and statistical interactions: an example from multistage carcinogenesis. *Int J Epidemiol*. 1981;10:383–387.
- Thompson WD. Effect modification and the limits of biological inference from epidemiologic data. *J Clin Epidemiol*. 1991;44:221–232.
- Vandenbroucke JP, von EE, Altman DG, et al. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *PLoS Med*. 2007;4:e297.
- Hosmer DW, Lemeshow S. Confidence interval estimation of interaction. *Epidemiology*. 1992;3:452–456.
- Knol MJ, van der Tweel I, Grobbee DE, et al. Estimating interaction on an additive scale between continuous determinants in a logistic regression model. *Int J Epidemiol*. 2007;36:1111–1118.
- Knol MJ, Vandenbroucke JP, Scott P, et al. What do case-control studies estimate? Survey of methods and assumptions in published case-control research. *Am J Epidemiol*. 15 September 2008. [Epub ahead of print].
- Garcia-Closas M, Malats N, Silverman D, et al. NAT2 slow acetylation, GSTM1 null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses. *Lancet*. 2005;366:649–659.
- Folsom AR, Cushman M, Heckbert SR, et al. Prospective study of fibrinolytic markers and venous thromboembolism. *J Clin Epidemiol*. 2003;56:598–603.
- Tavani A, Bertuzzi M, Gallus S, et al. Diabetes mellitus as a contributor to the risk of acute myocardial infarction. *J Clin Epidemiol*. 2002;55:1082–1087.
- Blom JW, Doggen CJ, Osanto S, et al. Malignancies, prothrombotic mutations, and the risk of venous thrombosis. *JAMA*. 2005;293:715–722.
- Lazo-Langner A, Knoll GA, Wells PS, et al. The risk of dialysis access thrombosis is related to the transforming growth factor-beta1 production haplotype and is modified by polymorphisms in the plasminogen activator inhibitor-type 1 gene. *Blood*. 2006;108:4052–4058.
- Schrager DL, Sinha R, Schroter S, et al. From submission to publication: a retrospective review of the tables and figures in a cohort of randomized controlled trials submitted to the British Medical Journal. *Ann Emerg Med*. 2006;48:750–756.
- Pocock SJ, Collier TJ, Dandreo KJ, et al. Issues in the reporting of epidemiological studies: a survey of recent practice. *BMJ*. 2004;329:883.
- Moreira ED Jr, Stein Z, Susser E. Reporting on methods of subgroup analysis in clinical trials: a survey of four scientific journals. *Braz J Med Biol Res*. 2001;34:1441–1446.
- Assmann SF, Pocock SJ, Enos LE, et al. Subgroup analysis and other (mis) uses of baseline data in clinical trials. *Lancet*. 2000;355:1064–1069.
- Rothwell PM. Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *Lancet*. 2005;365:176–186.
- Goodman SN. Multiple comparisons, explained. *Am J Epidemiol*. 1998;147:807–812.
- Rothman KJ. No adjustments are needed for multiple comparisons. *Epidemiology*. 1990;1:43–46.
- Szklo M, Nieto FJ. *Communicating Results of Epidemiologic Studies*. Epidemiology, Beyond the Basics. Sudbury, MA: Jones and Bartlett; 2000:chap 9.
- Greenland S. Additive risk versus additive relative risk models. *Epidemiology*. 1993;4:32–36.
- Skrondal A. Interaction as departure from additivity in case-control studies: a cautionary note. *Am J Epidemiol*. 2003;158:251–258.
- von Elm E, Altman DG, Egger M, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies [in Spanish]. *Ann Intern Med*. 2007;147:573–577.