

DAGitty

A Graphical Tool for Analyzing Causal Diagrams

To the Editor:

Causal diagrams, also known as directed acyclic graphs,^{1,2} provide an entirely graphical, yet mathematically rigorous methodology for minimizing bias in epidemiologic studies.^{3,4} The analysis of causal diagrams can be cumbersome in practice, and lends itself well to automatization by a computer program. Important first steps in this regard include the development of the DAG program by Knüppel and Stang⁵ and dagR by Breitling.⁶ We announce the release of DAGitty, which provides a graphical user interface tailored to draw and analyze causal diagrams. DAGitty overcomes some performance obstacles (pointed out by Breitling⁶) that affect earlier software when analyzing large diagrams.

The performance issues are 2-fold. First, previous software employed backtracking algorithms⁵ to enumerate and categorize all paths from exposure to outcome. This is a reasonable approach for small diagrams, but diagrams with tens of variables can already contain millions of paths. A full listing is of little interest to the human user, but can take hours or days to generate. Instead of a path list, DAGitty identifies the subdiagrams involved in causal and biasing paths and highlights them in different colors. This highlighting algorithm⁷ scales to very large diagrams. It provides a vivid impression about how causal and biasing effects “flow” in the diagram, that is, by which variables and causal arrows these effects are mediated.

The second problem with previous software has arisen when identifying minimally sufficient adjustment sets (MSA sets). According to causal diagram theory, adjustment for the covariates in an MSA set minimizes bias when estimating the total effect from exposure to outcome. A straightforward approach

to find MSA sets is to check each covariate set to see whether it is an MSA set. In a diagram with 50 covariates, this means that 2^{50} sets may have to be tested—a 16-digit number that is too large even for computers. To identify MSA sets more efficiently, we adapted an algorithm proposed recently for a related graph-theoretical problem.⁸ This algorithm is guaranteed to output the list of MSA sets reasonably quickly (ie, in polynomial time per MSA set output). Note, however, that very large or very regularly structured diagrams could in theory have millions of different MSA sets. If such diagrams become practically relevant, further research will be necessary to develop appropriate computational methods for helping the user to choose appropriate MSA sets.

The described algorithms enable DAGitty’s graphical interface to instantly reflect changes made to the diagram, such as adding a new arrow or inverting an arrow with unclear causal direction. This way, users can interactively assess the effects of their modifications on minimally sufficient adjustment sets and the flow of causal and biasing effects. We anticipate that these interactive possibilities will help users to develop an intuition about causal diagram theory, and to compare and decide among various causal diagrams.

DAGitty is available under an open-source license, allowing free access, redistribution, and modification. It runs out of the box in most modern web browsers and is available for online use and download at: www.dagitty.net.

ACKNOWLEDGMENTS

We thank Michael Elberfeld for discussions, and Sabine Schipf and Ines Polzer for providing example causal diagrams.

Johannes Textor

Institute for Theoretical Computer Science
University of Luebeck
Luebeck, Germany
textor@tcs.uni-luebeck.de

Juliane Hardt

Institute for Social Medicine
University of Luebeck
Luebeck, Germany

Sven Knüppel

Department of Epidemiology
German Institute of Human Nutrition
Potsdam-Rehbruecke
Nuthetal, Germany

REFERENCES

1. Pearl J. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press; 2000.
2. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*. 1999;10:37–48.
3. Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC Med Res Methodol*. 2008;8:70.
4. Glymour MM, Greenland S. Causal diagrams. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2008:183–209.
5. Knüppel S, Stang A. DAG program: identifying minimal sufficient adjustment sets [Letter]. *Epidemiology*. 2010;21:159.
6. Breitling L. dagR: a suite of R functions for directed acyclic graphs [Letter]. *Epidemiology*. 2010;21:586–587.
7. Textor J, Liškiewicz M. Adjustment criteria in causal diagrams: an algorithmic perspective. In: Cozman F, Pferrer A, eds. *Proceedings of the 27th Conference on Uncertainty in Artificial Intelligence*. Corvallis, OR: AUAI Press; in press.
8. Takata K. Space-optimal, backtracking algorithms to list the minimal vertex separators of a graph. *Discrete Applied Math*. 2010;158:1660–1667.

Transportability and Causal Generalization

To the Editor:

We read with interest the paper by Hernán and Vanderweele “On compound treatments and transportability of causal inference.”¹ They note 3 factors that influence the transportability (ie, generalizability) of the causal effect identified in a study—compound treatments, interference, and effect modification.

We think that consideration of transportability is long overdue. One limitation of the potential outcomes perspective is, in our view, a tendency to reify the causal effect estimated in a study as an actual intervention effect.² The relationship between the 2 needs explicit consideration, but the lack of direct correspondence does not necessarily im-

ply an “ill-defined” causal question, as suggested in much of the potential outcomes work in epidemiology.

Hernan and Vanderweele use new terminology to make points that have a long history in the Cook and Campbell tradition in psychology (eg, Shadish et al³). In that work, causal questions are divided into 2 broad categories: causal description and causal explanation. Causal description is about identifying what the causal effect was in a particular study. The focus is on internal validity. Causal explanation is about the transportability of effects. Shadish et al³ describe 2 issues in causal explanation: construct validity and external validity. Construct validity is the extent to which the treatment effect can be generalized beyond the particular operationalization in the study. In their framework, construct validity examines treatment variation irrelevance—the influence of various treatment versions on the causal effect. External validity is about the extent to which the treatment effect holds across various characteristics of persons and settings, addressing the issues of interference between units and effect modification.

The Cook and Campbell tradition shares with potential outcomes a counterfactual frame, although one starting from the perspective of Mackie rather than Lewis, who is favored in the potential outcomes literature. Shadish et al note the relationship between their view and the potential outcomes view (which they refer to as “Rubin’s causal model”):

“Rubin’s model is not intended to say much about the matters of casual generalization that we address in this book.”³ p.6

We have argued that integrating these aspects of the Cook and Campbell tradition into epidemiologic discussions of potential outcomes would be useful.^{2,4} We are glad to see a broader recognition of the problems of generalizability in the potential outcomes framework, which has been so illuminating for the explication of internal validity.

Sharon Schwartz

Department of Epidemiology
Columbia University
Mailman School of Public Health
New York, NY
sbs5@columbi.edu

Nicolle M. Gatto

Ulka B. Campbell

Department of Epidemiology
Columbia University
Mailman School of Public Health
New York, NY
Epidemiology
Worldwide Safety Strategy
Pfizer Inc.
New York, NY

REFERENCES

1. Hernan M, Vanderweele T. On compound treatments and transportability of causal inference. *Epidemiology*. 2011;22:368.
2. Schwartz S, Gatto N, Campbell U. What would have been is not what would be: counterfactuals of the past and potential outcomes of the future. In: Shroot PE, Keyes KM, Ornstein K, eds. *Causality and Psychopathology: Finding the Determinants of Disorders and Their Cures*. New York: Oxford; 2011:25–46.
3. Shadish WR, Cook TD, Campbell DT. *Experimental and Quasi-Experimental Designs for Generalized Casual Inference*. Boston: Houghton Mifflin; 2002.
4. Susser E, Schwartz S, Morabia A, Bromet EJ. *Psychiatric Epidemiology: Searching for the Causes of Mental Disorders*. New York: Oxford; 2006.

Forest Plot Viewer

A New Graphing Tool

To the Editor:

We would like to introduce the Forest Plot Viewer—a free program that creates figures to display point-estimate data with multiple columns of descriptive text that can be rapidly modified to display subsets of data based on filtering criteria. Figures in this format are called “forest plots” for reasons that remain somewhat obscure—perhaps because the figure summarizes a large amount of data that allows the reader to see the overall “forest” as well as the individual trees. This graphical format can be used to summarize results from several studies, or to present sets of results from a single study (for example, to show the effect of sequential adjustments¹).

The most common use of forest plots is to present data in systematic reviews or meta-analyses. Several software programs can generate such plots (eg, Comprehensive Meta-Analysis, RevMan, GraphPad, and R). In general, these programs are relatively inflexible, limiting the amount of accompanying text or requiring coding skills. The National Toxicology Program needed a more flexible and user-friendly interface to create figures with multiple columns of accompanying text. We developed such a tool (Forest Plot Viewer) in collaboration with SRA International. The text-based columns allow us to concisely present important study details that are not captured in point estimates (eg, study-population specifics, basis for disease diagnosis, exposure details, key adjustment factors) and to filter the display to include subsets of a larger data file.

Forest Plot Viewer horizontally plots a point estimate, confidence interval, and up to 15 columns of text for multiple rows of data (Figure). The shape, size, color, and fill of the point estimates can be modified to indicate groups or relative weights of results. One or 2 vertical reference lines can be included to indicate the deviation of the results from the null hypothesis or a meta-analysis result. A filtering function is available to select a subset of results for display from the larger set contained in the spreadsheet (eg, allowing the user to plot only prospective studies or only those with a particular exposure). The text and figures are automatically resized to fill the window space as settings are changed. Settings files can be saved for application to other datasets, or to create several displays from the same spreadsheet. Forest Plot Viewer does not perform a formal meta-analysis, but the results of a meta-analysis can be plotted with the program.

The program is freely available for download (http://ntp.niehs.nih.gov/go/tools_forestplotviewer) along with a user manual, example data, and example-settings files. Java version 5.0 or later is required to run the program (<http://www.java.com>). The data file is

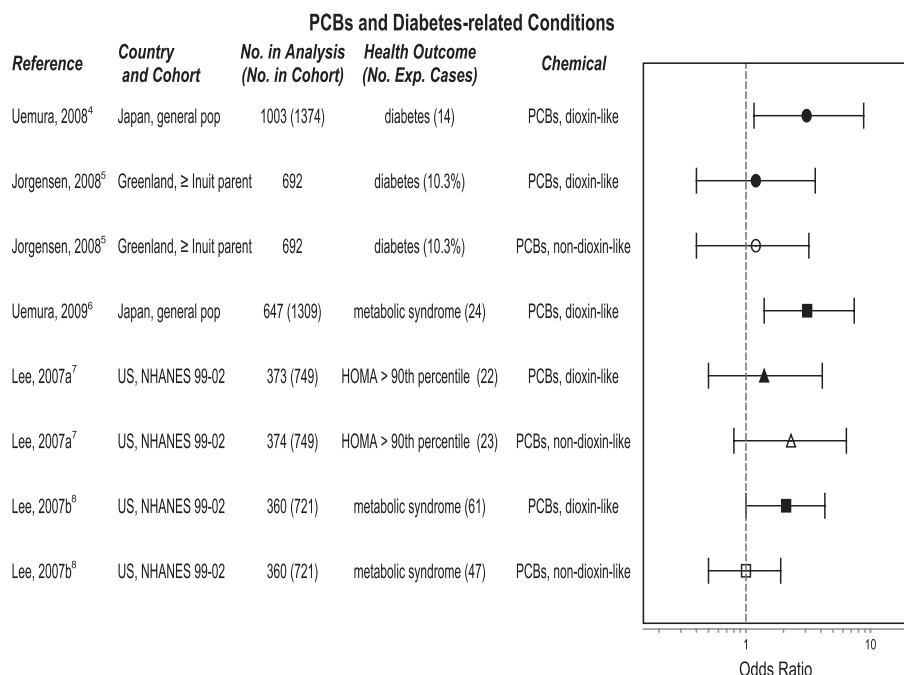


FIGURE. Studies of PCBs and diabetes-related conditions, including diabetes (circle), metabolic syndrome (square) and homeostatic model assessment (HOMA) of insulin resistance and beta cell function (triangle). Dioxin-like PCBs are indicated with filled shapes and non-dioxin-like PCBs have open shapes.

created by the user in a spreadsheet program such as Excel, or as a text file. Images from Forest Plot Viewer can be saved as standard image (GIF, PNG, BMP, JPG), scalable vector (SVG, EMF), or PDF files for use in presentations or publications.

The National Toxicology Program relied on this program during our January 2011 Workshop on the Role of Environmental Chemicals in the Development of Diabetes and Obesity to examine the role of persistent organic pollutants in diabetes and related health outcomes.² Experts who reviewed this literature found the software tool (and its ability to sort subcategories of studies easily during presentation) to be of great utility in evaluating a complex literature of almost 100 studies. One of the example datasets included is a subset of the studies considered during the workshop. The Figure displays studies of PCBs and diabetes related outcomes. Full references are available in the workshop materials.³

Abbe L. Boyles

Office of Health Assessment and Translation, National Toxicology Program
National Institute of Environmental Health Sciences
Durham, NC
boylesa@niehs.nih.gov

Shawn F. Harris

SRA International, Inc.
Durham, NC

Andrew A. Rooney

Kristina A. Thayer

Office of Health Assessment and Translation, National Toxicology Program
National Institute of Environmental Health Sciences
Durham, NC

REFERENCES

1. Baird DD, Travlos G, Wilson R, et al. Uterine leiomyomata in relation to insulin-like growth factor-I, insulin, and diabetes. *Epidemiology*. 2009;20:604–610.
2. NTP Workshop: Role of Environmental Chemicals in the Development of Diabetes and Obesity Breakout Group on Persistent Organic Pollutants (POPs). Available at: http://cerhr.niehs.nih.gov/evals/diabetesobesity/presentations/POPsFinal_508.pdf. Accessed May 18, 2011.
3. Appendix Table for Epidemiology Studies of Persistent Organic Pollutants (POPs). Available at:

http://cerhr.niehs.nih.gov/evals/diabetesobesity/presentations/POPsFinal_508.pdf. Accessed May 18, 2011.

4. Uemura H, Arisawa K, Hiyoshi M, et al. Associations of environmental exposure to dioxins with prevalent diabetes among general inhabitants in Japan. *Environ Res*. 2008;108:63–68.
5. Jorgensen ME, Borch-Johnsen K, Bjerregaard P. A cross-sectional study of the association between persistent organic pollutants and glucose intolerance among Greenland Inuit. *Diabetologia*. 2008;51:1416–1422.
6. Uemura H, Arisawa K, Hiyoshi M, et al. Prevalence of metabolic syndrome associated with body burden levels of dioxin and related compounds among Japan's general population. *Environ Health Perspect*. 2009;117:568–573.
7. Lee DH, Lee IK, Jin SH, Steffes M, Jacobs DR Jr. Association between serum concentrations of persistent organic pollutants and insulin resistance among nondiabetic adults: results from the National Health and Nutrition Examination Survey 1999–2002. *Diabetes Care*. 2007;30:622–628.
8. Lee DH, Lee IK, Porta M, Steffes M, Jacobs DR Jr. Relationship between serum concentrations of persistent organic pollutants and the prevalence of metabolic syndrome among non-diabetic adults: results from the National Health and Nutrition Examination Survey 1999–2002. *Diabetologia*. 2007;50:1841–1851.

Analgesics During Pregnancy and Undescended Testis

To the Editor:

Two recent papers suggested associations between the use of analgesics (paracetamol, aspirin, and ibuprofen) by women during pregnancy and the occurrence of congenital cryptorchidism in their offspring.^{1,2} This hypothesis deserves consideration given the widespread use of analgesics by pregnant women and the suspected secular increase in incidence of undescended testes in some areas.³ We therefore explored this association within a cohort of pregnant women in which specific information on the position of the testis has been collected at birth.

Women from Eden mother–child cohort were recruited before 24 gestational weeks in the maternity wards of Nancy and Poitiers (France) University hospitals, between 2003 and 2006.^{4,5} Exposure to analgesics during the first 2 trimesters of pregnancy was assessed from a questionnaire asked by the study

midwives between 24 and 28 gestational weeks. Women were asked “Since the beginning of the pregnancy, did you take one of the following drugs? Aspirin, paracetamol (listing 3 common preparations containing paracetamol) or ibuprofen (Advil).” Answers were recorded as yes/no/does not know, without distinguishing among the types of analgesic. A similar question was asked by the study midwife shortly after birth regarding analgesic use during the whole pregnancy. During a clinical examination conducted within 3 days after birth, pediatricians and study midwives assessed testis location (in scrotum, inguinal superficial, or not palpable). All singleton newborns with at least one testis not located in the scrotum were categorized as cases of undescended testes.

We studied the association between analgesics use and occurrence of undescended testes, using logistic regression adjusted for gestational duration (defined from the date of the last menstrual period^{5,6}), center, maternal age, parity, smoking, and educational level. Propensity-based methods provide an alternative to traditional adjustment to correct for confounding.⁷ Such methods may be particularly relevant when the number of cases (and hence the maximum number of covariates one can efficiently control for by adjustment⁸) is limited. We built a propensity score corresponding to the probability of use of analgesics as a function of maternal characteristics, and used this in additional weighted logistic models quantifying the impact of analgesics on occurrence of undescended testes using the inverse-probability-of-treatment weighted estimator.⁷

The frequency of undescended testis was 3.4% in Poitiers (16 cases, 465 births) and 5.1% in Nancy (22 cases, 430 births), giving an overall rate of 4.3% (4.1% after exclusion of preterm births). Analgesic use (overall frequency, 81%) was associated with parity ($P = 0.04$), toothache ($P = 0.02$), and antibiotic use ($P = 0.08$) during pregnancy. Frequency of undescended testis was 4.6% among offspring exposed to analgesics, compared

TABLE. Association of In Utero Exposure to Analgesics With Undescended Testis Among Male Singleton Live Births From Eden Cohort, 2003–2006

Analgesic Use	Total No.	Cases No. (%)	OR (95% CI)
Unadjusted models			
During 1st or 2nd trimester			
No ^a	172	5 (2.9)	1.0
Yes	723	33 (4.6)	1.6 (0.61–5.3)
During pregnancy			
No ^a	99	3 (3.0)	1.0
Yes	804	35 (4.4)	1.5 (0.45–7.5)
Models adjusted for center only			
During 1st or 2nd trimester			
No ^a	172	5 (2.9)	1.0
Yes	723	33 (4.6)	1.6 (0.59–4.0)
During pregnancy			
No ^a	99	3 (3.0)	1.0
Yes	804	35 (4.4)	1.4 (0.43–4.7)
Fully adjusted models^b			
During 1st or 2nd trimester			
No ^a	172	5 (2.9)	1.0
Yes	723	33 (4.6)	1.2 (0.45–3.2)
During pregnancy			
No ^a	99	3 (3.0)	1.0
Yes	804	35 (4.4)	1.1 (0.31–3.6)
IPTW approach based on propensity score^c			
During 1st or 2nd trimester			
No ^a	172	5 (2.9)	1.0
Yes	723	33 (4.6)	0.97 (0.35–2.7)
During pregnancy			
No ^a	99	3 (3.0)	1.0
Yes	804	35 (4.4)	0.72 (0.21–2.5)
Fully adjusted models,^b preterm births excluded			
During 1st or 2nd trimester			
No ^a	166	4 (2.4)	1.0
Yes	698	31 (4.4)	1.5 (0.51–4.4)
During pregnancy			
No ^a	96	2 (2.1)	1.0
Yes	776	33 (4.3)	1.6 (0.35–6.8)
IPTW approach based on propensity score,^c preterm births excluded			
During 1st or 2nd trimester			
No ^a	166	4 (2.4)	1.0
Yes	698	31 (4.4)	1.1 (0.36–3.4)
During pregnancy			
No ^a	96	2 (2.1)	1.0
Yes	776	33 (4.3)	0.88 (0.21–3.7)

^aReference category.

^bAdjusted for center, gestational duration (continuous term), maternal age (<25, 25–29, 30–34, ≥35 years), maternal smoking during second trimester (number of cigarettes, continuous term), maternal parity before the index birth (0, 1, 2, and more previous births), maternal education level.

^cWeighted logistic regression using an inverse probability density weighting based on a propensity score as defined by Kurth et al⁷ and adjusted for maternal smoking and gestational duration. The propensity score corresponded to the predicted probability of use of analgesics as a function of antibiotic use, anticoagulant use, tonics, hospitalization, toothache, hypertension, metrorrhagia, use of tranquilizer during pregnancy, parity, maternal age, level of education, center.

IPTW indicates inverse-probability-of-treatment weighted; OR, odds ratio.

with 2.9% among unexposed offspring (unadjusted odds ratio [OR] = 1.6 [95% confidence interval (CI) = 0.6–5.3]). After adjustment, the OR of undescended testis associated with use of analgesics during the 2 first trimesters of pregnancy was 1.2 (CI = 0.5–3.2; Table) in the whole population, and 1.5 among term births (CI = 0.5–4.4). ORs associated with analgesic use during the whole pregnancy were in the same range. The propensity-based weighting approach yielded ORs close to 1.0 (Table).

The main strength of our study is its longitudinal design. Limitations include small sample size, an inability to distinguish among specific compounds, and lack of information on dose, mixture, and exact timing of use. Our estimate of the association between maternal use of analgesic during the whole pregnancy and undescended testis risk (OR = 1.5 [95% CI = 0.5–7.5]; Table) was similar to that reported in the Danish population described by Kristensen et al¹ before adjustment (OR = 1.5 [95% CI = 0.8–2.8]), but weaker after adjustment (OR = 1.2 in our study vs. 1.4 in the Danish study). We adjusted for parity, as was done by Jensen et al² but not Kristensen et al. Our point estimates seemed sensitive to the choice of adjustment factors, to the exclusion of preterm births and propensity score weighting. Confidence intervals were broad, not allowing for conclusions regarding a possible impact of analgesics on the risk of undescended testes. We suggest that other cohorts with relevant exposure and outcome data report similar analyses, so that a meta-analysis can be performed.

Claire Philippat

Lise Giorgis-Allemand

Inserm, U823, IAB
Grenoble, France
Institut Albert Bonniot
Grenoble University
Grenoble, France

Cécile Chevrier

Sylvaine Cordier

Bernard Jégou

Inserm, U625, GERHM
Rennes, France
GERHM, IFR140
University of Rennes I
Rennes, France

Marie-Aline Charles

Inserm, U1018, CESP
Villejuif, France

Rémy Slama

Inserm, U823, IAB
Grenoble, France
Institut Albert Bonniot
Grenoble University
Grenoble, France
Remy.slama@ujf-grenoble.fr

REFERENCES

1. Kristensen DM, Hass U, Lesne L, et al. Intrauterine exposure to mild analgesics is a risk factor for development of male reproductive disorders in human and rat. *Hum Reprod*. 2011;26:235–244.
2. Jensen MS, Rebordosa C, Thulstrup AM, et al. Maternal use of acetaminophen, ibuprofen, and acetylsalicylic acid during pregnancy and risk of cryptorchidism. *Epidemiology*. 2010;21:779–785.
3. Boisen KA, Kaleva M, Main KM, et al. Difference in prevalence of congenital cryptorchidism in infants between two Nordic countries. *Lancet*. 2004;363:1264–1249.
4. Drouillet-Pinard P, Huel G, Slama R, et al. Prenatal mercury contamination: relationship with maternal seafood consumption during pregnancy and fetal growth in the 'EDEN mother-child' cohort. *Br J Nutr*. 2010;1–5.
5. Slama R, Thiebaugeorges O, Goua V, et al. Maternal personal exposure to airborne benzene and intrauterine growth. *Environ Health Perspect*. 2009;117:1313–1321.
6. Slama R, Khoshnood B, Kaminski M. How to control for gestational age in studies of effects of environmental factors on fetal growth? *Environ Health Perspect*. 2008;116:A284.
7. Kurth T, Walker AM, Glynn RJ, et al. Results of multivariable logistic regression, propensity matching, propensity adjustment, and propensity-based weighting under conditions of nonuniform effect. *Am J Epidemiol*. 2006;163:262–270.
8. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and cox regression. *Am J Epidemiol*. 2007;165:710–718.

Oral Disease and Risk of Pancreatic Cancer

To the Editor:

Pancreatic cancer is characterized by a largely unknown etiology and poor survival, stressing the need to identify

preventable risk factors. The few established risk factors (ie, tobacco smoking, chronic pancreatitis, diabetes, insulin resistance, and obesity) are only moderately or weakly associated with this tumor. A reported association between oral disease and pancreatic cancer remains to be confirmed in large and well-designed epidemiologic studies.^{1–5} Therefore, we conducted a large-scale case-control study, nested in a cohort comprising the nationwide Swedish Patient Register,⁶ during 1968–2008. Details of the study have been described elsewhere.⁷ All new cases of pancreatic cancer diagnosed during 1973–2008 with no prior cancer (except nonmelanoma skin cancer) were identified in the Swedish Cancer Register. Controls were frequency-matched using density-based sampling on the variables age, sex, calendar year, and year of the first record in the Patient Register. Hospitalization for a diagnosis of a defined disease of the oral cavity (ICD-8/9 codes 520–529 or ICD-10 codes K00–K14) since 1968, at least 5 years prior to inclusion, represented the study exposure.

Estimates of risk of pancreatic cancer were calculated as odds ratios (ORs) with 95% confidence intervals (CIs), using unconditional logistic regression analysis with multivariable adjustment for confounding. Data on potential confounding factors were collected from diagnoses recorded in the Patient Register: alcoholism or alcohol-related disease; chronic obstructive pulmonary disease or bronchitis; obesity; diabetes; gastric or duodenal ulcer; and pancreatitis.

The occurrence of disease of the oral cavity was slightly higher among the 29,218 pancreatic cancer cases than among the 99,992 control participants. The covariates pancreatitis, alcohol-related diseases, and diabetes were also overrepresented among cases compared with controls (data not shown). The adjusted odds ratio among patients with a history of oral disease was 1.2 (95% CI = 1.0–1.4) (Table).

These data provide support for a modest association between diseases

Supported by a grant from the Astrid and David Hagelén Foundation (to R.L.). The study was supported by grants from the Swedish Research Council (SIMSAM).

Copyright © 2011 by Lippincott Williams & Wilkins

ISSN: 1044-3983/11/2205-0749

DOI: 10.1097/EDE.0b013e318225bf48

TABLE. Disease of the Oral Cavity and the Risk of Pancreatic Cancer in Sweden During the Period 1973–2008

Adjusted for ^a	Pancreatic Cancer OR (95% CI)
Age and sex	1.2 (1.0–1.5)
Age, sex, alcohol, and chronic obstructive pulmonary disease	1.2 (1.0–1.5)
Age, sex, alcohol, chronic obstructive pulmonary disease, obesity, ulcer disease, diabetes, and pancreatitis	1.2 (1.0–1.4)

Chronic obstructive pulmonary disease or bronchitis; ICD-10 codes J40–J44 or ICD-8/9 codes 490–492.
 Obesity: ICD-10 codes E65 or E66, or ICD-9 code 278A.
 Diabetes: ICD-10 codes E10–E14, or ICD-8/9 code 250.
 Gastric or duodenal ulcer (*Helicobacter pylori* related disease); ICD-10 codes K25, K26, or K27, or ICD-8/9 codes 531, 532, or 533.
 Pancreatitis ICD-10 codes: K85, K86, or K87, or ICD-8/9 code 577.
^aAlcohol: ICD-10 codes K70, K85.2, F10, Y15, Y91 or ICD-8/9 codes 291, 303, 571.0, 571.2, 980.

of subsequent oral cavity and the risk of pancreatic cancer. Strengths include the large sample size, complete nationwide coverage of hospitalizations for oral cavity disease and of all cancers, and adjustments for confounding. Moreover, the nationwide register-based design removes possible recall or selection bias. By considering only oral disease that occurred at least 5 years before inclusion, we also reduce bias from reverse causation. Among weaknesses, only the most severe forms of oral cavity disease are captured because the exposure is based only on in-hospital care. Tobacco use is the most consistent risk factor for pancreatic cancer, and a main risk factor for periodontal disease.⁸ The use of hospitalization for tobacco-related diseases (chronic obstructive pulmonary disease and bronchitis) as surrogate assessment of tobacco use is crude, and leaves open the possibility of residual confounding. However, adjustment for known tobacco-related diseases had no effect on the estimate (Table).

The finding of an increased risk of pancreatic cancer after oral disease is supported by some previous research,^{3–5} whereas other studies have found no association.² While more research is needed to establish an association, possible biologic mechanisms include an increased production of nitrosamines due to an altered oral flora or systemic inflammation.

Rickard Ljung
 Upper Gastrointestinal Research
 Department of Molecular Medicine and
 Surgery
 Karolinska Institutet
 Stockholm, Sweden
 National Board of Health and Welfare
 Stockholm, Sweden
 rickard.ljung@ki.se

Tomas Sjöberg Bexelius
Lena Martin
 Upper Gastrointestinal Research
 Department of Molecular Medicine and
 Surgery
 Karolinska Institutet
 Stockholm, Sweden

Jesper Lagergren
 Upper Gastrointestinal Research
 Department of Molecular Medicine and
 Surgery
 Karolinska Institutet
 Stockholm, Sweden
 Division of Cancer Studies
 King's College
 London, United Kingdom

REFERENCES

1. Fitzpatrick SG, Katz J. The association between periodontal disease and cancer: a review of the literature. *J Dent.* 2010;38:83–95.
2. Hiraki A, Matsuo K, Suzuki T, Kawase T, Tajima K. Teeth loss and risk of cancer at 14 common sites in Japanese. *Cancer Epidemiol Biomarkers Prev.* 2008;17:1222–1227.
3. Hujoel PP, Drangsholt M, Spiekerman C, Weiss NS. An exploration of the periodontitis-cancer association. *Ann Epidemiol.* 2003;13:312–316.
4. Michaud DS, Joshupura K, Giovannucci E, Fuchs CS. A prospective study of periodontal disease and pancreatic cancer in US male health professionals. *J Natl Cancer Inst.* 2007;99:171–175.
5. Stolzenberg-Solomon RZ, Dodd KW, Blaser

MJ, Virtamo J, Taylor PR, Albanes D. Tooth loss, pancreatic cancer, and *Helicobacter pylori*. *Am J Clin Nutr.* 2003;78:176–181.

6. Rosen M. National Health Data Registers: a Nordic heritage to public health. *Scand J Public Health.* 2002;30:81–85.
7. Ljung R, Martin L, Lagergren J. Oral disease and risk of esophageal and gastric cancer in a nationwide nested case-control study in Sweden. *Eur J Cancer.* 2011. Doi:10.1016/j.ejca.2011.03.012.
8. Bergstrom J. Tobacco smoking and chronic destructive periodontal disease. *Odontology.* 2004;92:1–8.

Estimation of Population Percentiles

To the Editor:

Percentiles play an important part in descriptive statistics of continuous data, and their use is recommended for reference interval estimation.¹ We have selected various methods for the calculation of percentiles based on recommendations in the literature or use in popular software, and evaluated the accuracy of the percentile calculated in the sample as an estimate of the true population percentile,² using Monte-Carlo techniques.

All selected methods calculate a rank or an index that points to a number in the sorted array of sample data, and linear interpolation is applied when the index does not correspond to an integer value. One method (method A^{1,3}) calculates a rank or an index $p(n + 1)$ with p representing the centile (which is the percentile divided by 100) and n the sample size. Method B⁴ calculates an index $0.5 + pn$. Method C⁵ (commonly used in spreadsheets) uses $p(n - 1) + 1$ and method D⁶ uses

Supported by The National Institutes of Health, Bethesda, MD through a fellowship (to P.S.).
 Disclaimer: The views expressed do not necessarily represent the view of the National Institutes of Health, the Department of Health and Human Services, or the U.S. Federal Government.

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

Copyright © 2011 by Lippincott Williams & Wilkins
 ISSN: 1044-3983/11/2205-0750
 DOI: 10.1097/EDE.0b013e318225c1de

$p(n + 1/3) + 1/3$. Details of the use of these 4 methods are given in the eAppendix (<http://links.lww.com/EDE/A488>).

Experimental population data were obtained using a normal distribution pseudo-random number generator, programmed to generate a data set of 10^6 numbers with mean 0 and standard deviation 1.

From our population data, 6 sets of 100,000 random samples each were drawn using a pseudo-random number generator with uniform distribution. Each of these sets consisted of 100,000 random samples with sample size 20, 120, 500, and 1000. The average of the 5th and 95th percentiles obtained with the 4 methods in these sample sets were calculated and compared with the population values. For each sample, the relative difference with the population values was expressed as a percentage, and the mean and standard deviation of these percentages were calculated.

Next, the population data were transformed exponentially (base 10) to obtain a log-normal distribution and the experiments as described earlier were repeated.

The results for the 95th percentile in the normally distributed data are represented in the Table (more comprehensive tables with figures are available in the eAppendix, <http://links.lww.com/EDE/A488>). The results for the 5th percentile were symmetrical to the results for the 95th percentile and are not shown. Method B presents the highest accuracy, followed by method D, A, and C.

The results for the 5th and 95th percentile in the log-normally distributed data are represented in the Table. For the 5th percentile, method A has a higher accuracy than the methods D, B, and C, especially in small sample sizes, whereas for the 95th percentile method C presents the highest accuracy, followed by method B, D, and A.

We find that, for the calculation of percentiles, it may still be advantageous to transform log-normally distributed data. For example, the 95th percentile in the log-normal data should be about 44.16 ($10^{1.64}$). With method B and $n = 20$, we find an average of 114.69. But if we first

TABLE. Accuracy of Percentiles Calculated in Samples With Various Sizes From Normally and Log Normally Distributed Population Data

Methods	Sample Size	Normal Distribution 95th Percentile		Log-Normal Distribution			
		Average	% Difference (SD)	5th Percentile		95th Percentile	
				Average	% Difference (SD)	Average	% Difference (SD)
A: $p(n + 1)$	Population ^a	1.64		0.02		44.16	
	20	1.85	12.20 (31.06)	0.03	18.92 (141.23)	180.43	308.61 (1766.28)
	120	1.68	1.85 (11.88)	0.02	2.97 (46.85)	52.76	19.49 (59.86)
	500	1.65	0.47 (5.78)	0.02	0.82 (22.28)	46.01	4.20 (23.12)
B: $pn + 0.5$	1000	1.65	0.24 (4.07)	0.02	0.42 (15.69)	45.13	2.21 (16.00)
	Population ^a	1.64		0.02		44.16	
	20	1.64	-0.56 (25.48)	0.04	83.31 (179.54)	114.69	159.73 (973.07)
	120	1.64	-0.33 (11.37)	0.03	11.74 (48.61)	48.43	9.69 (52.26)
C: $p(n - 1) + 1$	500	1.64	-0.07 (5.72)	0.02	2.77 (22.48)	45.16	2.27 (22.52)
	1000	1.64	-0.02 (4.05)	0.02	1.44 (15.76)	44.64	1.09 (15.70)
	Population ^a	1.64		0.02		44.16	
	20	1.43	-12.97 (24.17)	0.06	147.81 (244.30)	48.07	8.87 (174.14)
D: $p(n + 1/3) + 1/3$	120	1.60	-2.50 (11.39)	0.03	20.77 (52.38)	44.18	0.05 (47.04)
	500	1.64	-0.59 (5.71)	0.02	4.82 (22.90)	44.24	0.19 (22.16)
	1000	1.64	-0.28 (4.06)	0.02	2.46 (15.92)	44.21	0.12 (15.60)
	Population ^a	1.64		0.02		44.16	
	20	1.71	3.80 (27.05)	0.04	61.98 (163.85)	138.42	213.47 (1367.82)
	120	1.65	0.44 (11.52)	0.02	8.55 (47.85)	49.96	13.15 (54.94)
	500	1.65	0.12 (5.69)	0.02	2.15 (22.37)	45.44	2.91 (22.72)
	1000	1.65	0.07 (4.07)	0.02	1.07 (15.72)	44.79	1.43 (15.79)

^a $n = 10^6$.

Results are presented as average percentile, the average of the relative differences (%), and their standard deviation (SD).

log-transform the data we find, on average, 1.64, which back-transforms to $10^{1.64}$ or 43.65, which is much closer to the true population value of 44.16. The effect of the log-transformation may be explained by the fact that linear interpolation is applied in the calculations of percentiles, and the transformation changes the distribution model within the interpolated interval.

We conclude that method B is the preferred method in general for continuous data, taking into account the recommendation to transform the data to a normal distribution if necessary.

Finally, the large standard deviations of the observed differences illustrate the large statistical uncertainty associated with the estimated percentiles in small sample sizes. Therefore, we stress the importance of reporting percentiles with their 95% confidence interval.

Frank Schoonjans

Department of Endocrinology
Ghent University Hospital
Gent, Belgium
MedCalc Software bvba
Mariakerke, Belgium
frank.schoonjans@ugent.be

Dirk De Bacquer

Department of Public Health
Ghent University
Gent, Belgium

Pirmin Schmid

Department of Transfusion Medicine
Clinical Center
National Institutes of Health
Bethesda, MD

REFERENCES

1. CLSI. *Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory: Approved Guideline*. 3rd ed. Wayne, PA: Clinical and Laboratory Standards Institute; 2008. CLSI Document C28-A3.
2. Walter S. Problems with percentiles. *Int J Epidemiol*. 1986;15:431-532.
3. Altman DG. *Practical Statistics for Medical Research*. London: Chapman and Hall; 1991.
4. Armitage P, Berry G, Matthews JN. *Statistical Methods in Medical Research*. 4th ed. Oxford: Blackwell Science; 2002.
5. Gumbel EJ. La Probabilité des Hypothèses. *Comptes Rendus de l'Académie des Sciences (Paris)*. 1939;209:645-647.
6. Hyndman RJ, Fan Y. Sample quantiles in statistical packages. *Am Stat*. 1996;50:361-365.