The increasing sophistication of statistics in the New England Journal of Medicine

Nicholas J. Horton

Smith College and University of Auckland

April 17, 2008, ASA TSHS webinar

nhorton@email.smith.edu http://www.math.smith.edu/~nhorton

・ 同 ト ・ ヨ ト ・ ヨ ト

or what your physician should know about statistics (but perhaps doesn't)

Nicholas J. Horton

Smith College and University of Auckland

April 17, 2008, ASA TSHS webinar

nhorton@email.smith.edu http://www.math.smith.edu/~nhorton

- 4 同 6 4 日 6 4 日 6

Acknowledgements

- joint work with Suzanne Switzer (SC 2006)
- partial funding from Howard Hughes Medical Institute
- published letter in the *New England Journal of Medicine* (2005) and article in *CHANCE* (2007)

・ 戸 ト ・ ヨ ト ・ ヨ

Introduction and motivation

- increasing sophistication in statistical methods
- need for creative ways to ensure citizens are quantitatively savvy
- implications for educators at all levels
- case study: statistics needed by physicians
- but first, a warmup exercise and cautionary note

・ 戸 ト ・ ヨ ト ・ ヨ

Warm up exercise Cautionary note: side effects of Vioxx

Warm up exercise (Windish et al, JAMA, 2007)

The Third National Health and Nutrition Examination Survey was conducted in the United States in the 1990's to examine the relationship between obesity and depression. The authors investigated the association between major depression and body mass index (BMI) for males and females.

・ 同 ト ・ ヨ ト ・ ヨ ト

Warm up exercise Cautionary note: side effects of Vioxx

Table of results

Table. Unadjusted Odds Ratios of Major Depression During the Past Month

BMI Category, kg/m ²	Unadjusted Odds Ratio	95% Confidence Interval
Normal weight (BMI 18.5-24.9)	1.00	
Underweight (BMI < 18.5)	1.17	0.49-2.80
Overweight (BMI 25.0-29.9)	0.86	0.53-1.41
Obese (BMI ≥ 30)	1.88	1.02-3.46
Class 1 (BMI 30-34.9)	1.28	0.64-2.56
Class 2 (BMI 35-39.9)	1.76	0.78-3.95
Class 3 (BMI ≥ 40)	4.98	2.07-11.99

The increasing sophistication of statistics in the NEJM

Nicholas J. Horton

Warm up exercise Cautionary note: side effects of Vioxx

Warm up exercise

From the table, what is the correct interpretation of the overweight values of 0.86?

- Overweight individuals' odds of having major depression are 14% lower than the odds of having major depression for individuals with normal weight
- Overweight individuals' odds of having major depression are 14% higher than the odds of having major depression for individuals with normal weight
- c. An overweight individual has a 0.86 probability of having major depression
- d. An overweight individual has a 0.86 odds of having major depression

- 4 同 ト 4 ヨ ト 4 ヨ ト

Warm up exercise Cautionary note: side effects of Vioxx

Warm up exercise

From the table, what is the correct interpretation of the overweight values of 0.86?

- Overweight individuals' odds of having major depression are 14% lower than the odds of having major depression for individuals with normal weight
- Overweight individuals' odds of having major depression are 14% higher than the odds of having major depression for individuals with normal weight
- c. An overweight individual has a 0.86 probability of having major depression
- d. An overweight individual has a 0.86 odds of having major depression

39% correct (95% CI from 33-45%) in a sample of Connecticut area residents

Warm up exercise Cautionary note: side effects of Vioxx

A brief history of Vioxx

- approved by FDA in 1999
- used by more than 80 million people
- annual sales revenue of \$2.5B
- withdrawn from market (voluntarily by Merck) in Fall 2004 because of concerns about heart attack risk
- estimates of excess heart attacks due to Vioxx range from 88,000-139,000 (30-40% fatal)

- 4 同 2 4 日 2 4 日 2

Warm up exercise Cautionary note: side effects of Vioxx

Bombardier et al (NEJM 2000)

Results: Rofecoxib (Vioxx) and naproxen had similar efficacy against rheumatoid arthritis. During a median follow-up of 9.0 months, 2.1 confirmed gastrointestinal events per 100 patient-years occurred with Vioxx, as compared with 4.5 per 100 patient-years with naproxen (relative risk, 0.5; 95 percent confidence interval, 0.3 to 0.6; P < 0.001). The respective rates of complicated confirmed events (perforation, obstruction, and severe upper gastrointestinal bleeding) were 0.6 per 100 patient-years and 1.4 per 100 patient-years (relative risk, 0.4; 95 percent confidence interval, 0.2 to 0.8; P=0.005).

< ロ > < 同 > < 回 > < 回 >

Warm up exercise Cautionary note: side effects of Vioxx

Bombardier et al (NEJM 2000)

The incidence of myocardial infarction was lower among patients in the naproxen group than among those in the Vioxx group (0.1 percent vs. 0.4 percent; relative risk, 0.2; 95 percent confidence interval, 0.1 to 0.7); the overall mortality rate and the rate of death from cardiovascular causes were similar in the two groups.

(4月) (1日) (1日)

Warm up exercise Cautionary note: side effects of Vioxx

Bombardier et al (NEJM 2000)



Warm up exercise Cautionary note: side effects of Vioxx

Bombardier et al (NEJM 2000)



(日) (同) (三) (三)

Warm up exercise Cautionary note: side effects of Vioxx

Bombardier et al (NEJM 2000)



Warm up exercise Cautionary note: side effects of Vioxx

Bombardier et al (NEJM 2000)



• distance from 0.2 to 1 is 0.8, while distance from 1 to 5 is 4

Warm up exercise Cautionary note: side effects of Vioxx

Bombardier et al (NEJM 2000)



- distance from 0.2 to 1 is 0.8, while distance from 1 to 5 is 4
- RR of 0.2 (comparing naproxen to Vioxx) is equivalent to RR of 5 (comparing Vioxx to naproxen)

A (1) > (1) = (1) (1)

Warm up exercise Cautionary note: side effects of Vioxx

Bombardier et al (NEJM 2000)



- distance from 0.2 to 1 is 0.8, while distance from 1 to 5 is 4
- RR of 0.2 (comparing naproxen to Vioxx) is equivalent to RR of 5 (comparing Vioxx to naproxen)
- article appeared in 2000
- not taken off market until 2004
- was the RR of 0.2 well understood?

A (10) < A (10) </p>

Introduction Why the NEJM? Methods Results



- bolster my claim that statistical methods are becoming more sophisticated
- assess changes in complexity of statistics used in research reports published in a widely read medical journal
- consider the implications for statistical education of clinicians (as well as everyone else!)

Introduction Why the NEJM? Methods Results

Background

- prior work of Emerson and Colditz described the use of statistics in the New England Journal of Medicine (NEJM) for 1978-1979 and 1989
- updated their studies to see what statistical techniques are used 15 years later in Original Articles published in NEJM
- used similar methodology to facilitate comparisons

- 4 同 2 4 日 2 4 日 2

Introduction Why the NEJM? Methods Results



- historical journal of record
- ostensibly read by clinicians (not just researcher types)
- widely disseminated
- high impact

(日) (同) (三) (三)

Introduction Why the NEJM? Methods Results

History of the NEJM

- founded as a quarterly in 1812
- became a weekly publication in 1828
- one early editor resigned to become mayor of Boston
- now one of the core medical journals (along with the Lancet, JAMA, BMJ)

- 4 同 2 4 日 2 4 日

Introduction Why the NEJM? Methods Results

Dissemination

- NEJM is to internal medicine what Nature and Science are to general science (Wikipedia)
- all contents available freely online after 6 months (all articles freely available to readers from economically disadvantaged countries)
- ostensibly read by clinicians (not just researcher types), large circulation
- audio version downloadable for iPods
- impact factor for 2004 = 38.57!
- widely read journal with impact

▲ □ ▶ ▲ □ ▶ ▲ □ ▶

Introduction Why the NEJM? Methods Results

Prior work of Emerson and Colditz

- Volumes 298 through 301 (January 1978 through December 1979), reviewed 332 original articles
- Volume 321 (July through December 1989), reviewed 115 original articles
- used arbitrary but useful hierarchy

- 4 同 ト 4 ヨ ト 4 ヨ ト

Introduction Why the NEJM? Methods Results

Emerson and Colditz hierarchy

- no statistics (or purely descriptive)
- t-tests
- contingency tables
- nonparametric tests
- epidemiological statistics
- Pearson correlation/simple linear regression
- analysis of variance
- transformations

・ 同 ト ・ ヨ ト ・ ヨ

Introduction Why the NEJM? Methods Results

Emerson and Colditz hierarchy (cont.)

- survival methods
- multiple regression
- multiple comparisons
- multiway tables
- opower

▲ □ ▶ ▲ □ ▶ ▲

Introduction Why the NEJM? Methods Results

Categories we added

- sensitivity analysis
- repeated measures analysis
- missing data methods
- noninferiority trials
- receiver operating characteristic curves
- resampling based inference
- principal components analysis/cluster analysis
- other methods (genetics, meta-analysis, etc.)

- 4 同 ト 4 ヨ ト 4 ヨ ト

Introduction Why the NEJM? Methods Results



- independent coding by both authors
- consensus review process when reports were discordant
- time-consuming process (even for only a few hundred papers)

The NEW ENGLAND JOURNAL of MEDICINE

HOME	SEARCH CURRENT	issue F	PAST ISSUES	COLLECTIONS	HELP
	TAE	BLE OF C	ONTENTS		
<u> </u>	Volume 350	June 24,	2004	Number 26	<u>Next</u> ►
THIS WEEK IN THE JO	URNAL		PERSPECTIVE		
Article Summaries			Small Airway	's in COPD	
Summaries PDF			P. J. Barnes		
			Extract Ful	I Text PDF	
ORIGINAL ARTICLES	A investigation in		D II (' D		
Chronic Obstructive P	-Alfway Obstruction in Pulmonary Disease		Pediatric Ken	ial -Replaceme	ent Therapy —
J. C. Hogg and Other	rs		D S Mill ine	je r	
Abstract Full Text	PDF		Extract Ful	I Text I PDF	
	·				
Long -Term Survival o	of Children with End -	-Stage	Milk of Hum	an Kindness?	— HAMLET, Human
Renal Disease			Papillomavir	us, and Warts	
S. P. MCDONAIG and C	J. C. Claig		J. N. Bouwes	Bavinck and N	M. C.W. Feltkamp
			Extract Ful	I Text PDF	
Treatment of Skin Pan	oillomas with Topical	α -	Dowowful Con	nog Myggtat	in Domilation of
Lactalbumin –Oleic A	Acid		Human Musc	les — Myostau de Mass	in Regulation of
L. Gustafsson and Ot	hers		E. M. McNall	V	
Abstract Full Text	PDF CME Exam		Extract Ful	Text PDF	
Folato Thomany and In	Stant Destances off	0H			
Coronary Stanting	-Stellt Restellosis alt	er			
H Lange and Others			EDITORIALS		
Abstract Full Text	PDF		Prevention of	Cardiovasculai	r Events after
			H C Horme	Coronary Intel	rvenuo n
Brief Report: Myostati	n Mutation Associated		Extract Ful		
with Gross Muscle Hy	pertrophy in a Child				

M. Schuelke and Others Extract | Full Text | PDF



	HOME	SEARCH CURREN	NT ISSUE PAST ISSUE	S COLLECTIONS HEL	P
		т	ABLE OF CONTENT	rs	
	Previou s	Volume 350	June 24, 2004	Number 26	<u>Next</u> >
	THIS WEEK IN THE JOUR Article Summaries Summaries PDF	NAL	PERSPECT Small Ai P. J. Bai Extract	IVE rways in COPD nes Full Text PDF	
Folate The H. Lange a <u>Abstract</u>]	rapy and In -S and Others Full Text PD	tent Reste <u>F</u>	enosis after	· Coronary S	tenting
					amp
	Treatme Lactalbumin – L. Gustafsson and Oue <u>Abstract</u> <u>Full Text</u> <u>P</u>	DF 12	L. M. Me Extract	Myostatin Reg Muscle Mass Nally Full Text PDF	ulation of
	Folate Therapy and In Coronary Stenting H. Lange and Others <u>Abstract</u> <u>Full Text</u> <u>P</u>	-Stent Restenosis aft	er EDITOR Preventi	IALS on of Cardiovascular Even	ts after

Brief Report: Myostatin Mutation Associated with Gross Muscle Hypertrophy in a Child M. Schuelke and Others Extract | Full Text | PDF Prevention of Cardiovascular Events afte Percutaneous Coronary Interventio n H. C. Hermann <u>Extract</u> | <u>Full Text</u> | <u>PDF</u>

	ORIGINA	AL ARTICLE		
Previous	Volume 350:2673-2681	<u>June 24, 2004</u>	Number 26	<u>Next</u> ►

Folate Therapy and In-Stent Restenosis after Coronary Stenting

Helmut Lange, M.D., Harry Suryapranata, M.D., Giuseppe De Luca, M.D., Caspar Börner, M.D., Joep Dille, B.Sc., Klaus Kallmayer, M.D.,

ABSTRACT

Background Vitamin therapy to lower homocysteine levels has recently been recommended for the prevention of restenosis after coronary angioplasty. We tested the effect of a combination of folic acid, vitamin B_{6} , and vitamin B_{12} (referred to as folate therapy) on the risk of angiographic restenosis after coronary-stent placement in a double-blind, multicenter trial.

Methods A total of 636 patients who had undergone successful coronary stenting were randomly assigned to receive 1 mg of folic acid, 5 mg of vitamin B_6 , and 1 mg of vitamin B_{12} intravenously, followed by daily oral doses of 1.2 mg of folic acid, 48 mg of vitamin B_6 , and 60 µg of vitamin B_{12} for six months, or to receive placebo. The angiographic end points (minimal luminal diameter, late loss, and restenosis rate) were assessed at six months by means of quantitative coronary angiography.

Results At follow-up, the mean (\pm SD) minimal luminal diameter was significantly smaller in the folate group than in the placebo group (1.59 \pm 0.62 mm vs. 1.74 \pm 0.64 mm, P=0.008), and the extent of late luminal loss was greater (0.90 \pm 0.55 mm vs. 0.76 \pm 0.58 mm, P=0.004). The restenosis rate was higher in the folate group than in the placebo group (34.5 percent vs. 26.5 percent, P=0.05), and a higher percentage of patients in the folate group required repeated target-vessel revascularization (15.8 percent vs. 10.6 percent, P=0.05). Folate therapy had adverse effects on the risk of restenosis in all subgroups except for women, patients with diabetes, and patients with markedly elevated homocysteine levels (15 µmol per liter or more) at baseline.

Conclusions Contrary to previous findings, the administration of folate, vitamin B_6 , and vitamin B_{12} after coronary stenting may increase the risk of in-stent restenosis and the need for target-vessel revascularization.

Statistical Analysis

We estimated that 622 patients would need to be enrolled for the study to have the statistical power to detect a difference in late luminal loss of 0.13 mm between the two groups with 90 percent power, assuming a standard deviation of 0.50 mm in each group, using a two-group t-test and a two-sided significance level of 0.05. To account for the probability that some patients would not complete the study protocol, we planned to enroll 650 patients. **Power** : use of the size of detectable (or useful) difference in determining sample size

Statistical Analysis

We estimated that 622 patients would need to be enrolled for the study to have the statistical power to detect a difference in late luminal loss of 0.13 mm between the two groups with 90 percent power, assuming a standard deviation of 0.50 mm in each group, using a two-group t-test and a two-sided significance level of 0.05. To account for the probability that some patients would not complete the study protocol, we planned to enroll 650 patients.

t-test : One-sample, matched-pair and two-sample t-tests



Contingency tables : Chi-square tests, Fisher's exact test, McNemar's test

Multiway tables : Mantel-Haenszel procedure, log-linear models

Regression for survival : Logistic

regression (formerly included Cox regression)

Statistical Analysis (cont.)

A multiple logistic-regression analysis was performed to identify independent variables associated with restenosis. The stepwise selection of the variable and the estimation of significant probabilities were performed by means of the maximum-likelihood ratio test. The chi-square value was calculated from the log of the ratio of maximum-partial-likelihood functions. The additional value of each category of variables added sequentially was evaluated on the basis of the increases in the overall likelihood statistic ratic. The following variables, if significant on univariate analysis, were included in the multivariate analysis: age; sex; the presence or absence of diabetes (defined by current use of insulin or oral hypoglycemic therapy), shoking, previous myocardial infarction, previous bypass surgery, and use of plycoprotein I b/IIIa inhibitors; baseline and follow-up levels of homocysteine, cholesterol, and triglycerides; lesion length; reference diameter; postprocedural minimar luminal diameter; and type

of therapy at discharge (statins, beta enzyme inhibitors).

Multiple Regression : Includes polynomial regression and stepwise regression, smoothing, splines, generalized additive models **Life table :** Actuarial life table, Kaplan-Meier estimate of survival

Statistical Analysis (cont.)

Adverse events during follow-up were analyzed by the Kaplan–Meier method. Differences in the event-free survival curves between the two groups were compared with the use of the log-rank test. Data analysis was performed by an independent core laboratory (Diagram). The investigators initiated the study, wrote the article, and had full access to the data.

Other survival analysis : Breslow's extension of Kruskal-Wallis, log rank, Cox proportional hazards models

Results

Angiographic Analysis

Epidemiologic statistics : Relative risk, odds ratio, log odds, measures of association, sensitivity, specificity, excess risk, incidence density, kappa statistic (agreement)

We performed subgroup analyses of the relative risk of restends is in order to evaluate whether the increase in restenosis was distributed evenly across subgroups (Figure 2). The point estimates showed that folate therapy reduced the risk of restenosis, but not significantly so, among women (relative risk, 0.67; 95 percent confidence interval, 0.39 to 1.14; P=0.13; P for heterogeneity = 0.002), and among patients with diabetes, as compared with patients without diabetes (relative risk, 0.71; 95 percent confidence interval, 0.44 to 1.15; P=0.16; P for heterogeneity = 0.02). Likewise, patients with baseline homocysteine levels of 15 µmol per liter or more had a decreased risk of restenosis with the use of folate therapy (relative risk, 0.86; 95 percent confidence interval, 0.49 to 1.52; P=0.61), whereas those with lower levels of homocysteine had an increase in risk (relative risk, 1.42; 95 percent confidence interval, 1.05 to 1.92; P=0.02; P for heterogeneity = 0.12). The adverse effect of folate therapy was most pronounced in patients who had vessels with a reference diameter of 3 mm or more (relative risk, 2.02; 95 percent confidence interval, 1.18 to 3.48; P=0.008).

Nick's Data Collection Sheet

rolume/No./15 : E 350/26 Sur 24, 2004 <u>Author/Article</u>	No Statistical Methods	Contingency Tables	Multiway Tables	Epidemiologic statistics	t-tests	Pearson correlation	Simple linear regression	Multiple regression	Analysis of Variance	Multiple Comparisons	Nonparametric tests	Nonparametric correlation	Life table	Non-Linear or GLM	Other survival analysis	Adjustment and Standardization	Sensitivity analysis	Power	Transformation	Cost-benefit Analysis	Missing data methods	Bayesian Analysis	Repeated Measures/ Longitudinal analysis	Structural equation model	Receiver operating Characteristic	-
hoggetal						12	\checkmark	\checkmark				1														10:50-
"long term surreal" Accounted et al	1			1				/	. 0		1		7		\checkmark											_
"treation t" gustefssin et al]			1																					
"Folate Tlerapy" boyse et al]	V Breston day	\checkmark	1			/	D]		\checkmark	1	/			\checkmark								-
"BR: mjostatin" schelke	/																-									
	1	1		1	1																					-

Nick's Data Collection Sheet

											r(•	(0-		_
olume/No./Is_ 3																ation							dinal		ristic	
\$ 350/26	ds			tics			ion			SU		ation			sis	lardiz				s	s		ongitue	lodel	aracte	
Jun 24 ZONY	Metho	ables	es	statis		ation	egress	ssion	riance	ariso	tests	correl		GLM	analy	Stand	lysis		-	nalysi	lethod	ysis	res/ Lo	tion n	ng Ch	
, ,	tical N	ncy T	Tabl	logic		orrela	lear r	regree	of Va	Comp	netric	netric		ar or	vival	nt and	y ana		natior	efit A	lata m	Anal	Measu	equa	perati	
	Statis	itinge	ltiway	demio	sts	rson c	ple lir	ltiple	lysis	ltiple	iparar	paran	table	h-Line	er sur	ustme	sitivit	/er	nsforr	t-bene	sing c	esian	cated l	ctural	eiver o	
Author/Article	No	Con	Mul	Epi	t-tes	Pear	Sim	Mul	Ana	Mul	Nor	Non	Life	Nor	Oth	Adji	Sen	Pow	Tra	Cos	Mis	Bay	Rep	Stru	Reco	
"Folate Therapy"			,						1																	
large et al			\checkmark	1	1			1	VA		1		1	1	1			\checkmark								
			Briston	V	V			\checkmark	00		5		\vee	V	V	_		Y								
			indy																		-					_

"Folate Therapy" Loge et al J V No VV V V day "BR: myo statin" schelke													
nBR: mjostatin ^h schelke	Folate Tlerapy" boye et al]	J Breston J day	\checkmark	$\overline{\mathbf{A}}$	De		VV	\bigvee	 \checkmark			
	"BR: myostatin" schelke	/										, I	

Suzi's Data Collection Sheet

Volume/No./Issue 350 26 Jane 24 2004	Vo Statistical Methods	Contingency Tables	Aultiway Tables	pidemiologic statistics	tests	earson correlation	imple linear regression	Aultiple regression	unalysis of Variance	Aultiple Comparisons	Ionparametric tests	lonparametric correlation	ife table	Ion-Linear or GLM	other survival analysis	djustment and Standardization	ensitivity analysis	ower	ransformation	ost-benefīt Analysis	fissing data methods	ayesian Analysis	epeated Measures/ Longitudinal	tructural equation model	eceiver operating Characteristic	ther
Author/Article	4	0	4	Ш	4	Р	S	N	A	Z	Z	Z	Г	Z	0	A	S	P	H	0	2	B	R	S	R	0
Lange et al																										KA
Folate Therarpy		\bigvee	1		V.			V	1		\checkmark		V	J	N			\checkmark								



ORIGINAL ARTICLE

Previous Volume 351:552-559 <u>August 5, 2004</u> Number 6 <u>Next</u>

Loss of Smad3 in Acute T-Cell Lymphoblastic Leukemia

Lawrence A. Wolfraim, Ph.D., Tania M. Fernandez, Ph.D., Mizuko Mamura, M.D., Ph.D., Walter L. Fuller, B.S., Rajesh Kumar, Ph.D., Diane E. Cole, B.S., Stacey Byfield, Ph.D., Angelina Felici, Ph.D., Kathleen C. Flanders, Ph.D., Thomas M. Walz, M.D., Anita B. Roberts, Ph.D., Peter D. Aplan, M.D., Frank M. Balis, M.D., and John J. Letterio, M.D.

- Methods
 - Patients and Cell Collection
 - Western Blotting
 - Determination of Free and Cyclin D3-Associated p27Kip1
 - Analysis of Smad3 Messenger RNA
 - Sequencing of the MADH3 Gene
 - Proliferation Assays and Interleukin-2 Measurements
 - Immunohistochemistry
- Results
 - Loss of Smad3 in T-Cell ALL
 - Levels of Smad3 and T-Cell Responses to TGF- in Mice
 - Smad3 and p27Kip1 in the Suppression of Leukemogenesis
- Discussion

Sequencing of the MADH3 Gene

To prepare genomic DNA, leukemia cells were digested in 300 µl of lysis buffer (200 mM sodium chloride; 40 mM TRIS, pH 8.0; 20 mM EDTA; 0.5 percent sodium dodecyl sulfate; 0.5 percent ^βmercaptoethanol; and 2 mg of fresh proteinase K per milliliter), and DNA was precipitated from lysates by the addition of 600 µl of ice-cold ethanol, then rinsed, and dissolved in water. The following primers were used for PCR amplification and sequencing of human MADH3 exons: exon 1, 5'CGAAGTTTGGGCGACCGCGG3' (forward) and 5'CTCTCTCCCTCTTCCCATCTCCAGC3' (reverse), yielding a 420-bp PCR fragment; exon 2, 5'Caatcacatttccctctttctg3' (forward) and 5'Cagcatacctggtgtctctac3' (reverse), yielding a 246-bp fragment; exon 3, 5'Gtctttgcaaaaggtgtctc3' (forward) and 5'Ctgctaatcagttaagaataag3' (reverse), yielding a 190-bp product; exon 4, 5'Caggccaagaatcttttgtgaag3' (forward) and 5'Aaacctggcatatggttgtctttc3' (reverse), yielding a 319-bp product; exon 5, 5'Gagattataatccctctgaaatgc3' (forward) and 5'Ctgcattcctgttgacattg3' (reverse), yielding a 305-bp product; exon 6, 5'Cattgtgtgtgagcaaag3' (forward) and 5'Cacctccagattgacaacgcaatc3' (reverse), yielding a 315-bp product; exon 7, 5'Ctgtttctgtgtttttggcag3' (forward) and Cttggcctctctctgatctttg3' (reverse), yielding a 306-bp product; exon 8, bp product; and exon 9, 5'GGTAGAGGAGTTTGGCCGGGTAGTT3' (forward) and 5'TGGGGCCAAAGGGTAAATGTGTT3' (reverse), yielding a 521-bp product.

Introduction Why the NEJM? Methods Results

Results

Procedure	78-79	1989	04-05	Accumulated by
No methods or descriptive	27%	12%	13%	13%
t-tests	44%	39%	26%	14%
contingency tables	27%	36%	53%	15%
nonparametric tests	11%	21%	27%	17%
epidemiologic statistics	9%	22%	35%	18%
simple linear regression	8%	9%	6%	18%
analysis of variance	8%	20%	16%	20%
transformation	7%	7%	10%	20%

<ロ> <部> < 2> < 2> < 2> < 2> < 2</p>

э.

Introduction Why the NEJM? Methods Results

Results

Procedure	78-79	1989	04-05	Accumulated by
survival methods	11%	32%	61%	24%
multiple regression	5%	14%	51%	39%
multiple comparisons	3%	9%	23%	41%
multiway tables	4%	10%	13%	44%
power	3%	3%	39%	68%
sensitivity analysis	0%	0%	6%	72%
repeated measures	-	-	12%	80%
missing data methods	-	-	8%	87%
noninferiority trials	-	-	4%	91%
ROC	-	-	2%	93%
resampling	-	-	2%	94%
other	-	-	6%	100%

Nicholas J. Horton The increasing sophistication of statistics in the NEJM

<ロ> <部> < 2> < 2> < 2> < 2> < 2</p>

э.

Introduction Why the NEJM? Methods Results

Graphical display of Accumulated by metric



Nicholas J. Horton The increasing sophistication of statistics in the NEJM

Introduction Why the NEJM? Methods Results

Average method uses per article

Year	1978-1979	1989	2004-2005
Average	1.9	2.7	4.2

(日) (同) (三) (三)

Introduction Why the NEJM? Methods Results

Results

- increase in the number of methods used
- topics typically included in introductory statistics courses make 21% of the articles "accessible" in the sense of Emerson and Colditz
- the biggest jumps in *Accumulated by* percentage related to knowledge of multiple regression, power, and repeated measures/missing data.

< ロ > < 同 > < 回 > < 回 >

Introduction Why the NEJM? Methods Results

Limitations

- arbitrary coding of hierarchy that does not typically correspond to the ordering that topics are presented in textbooks and courses
- changes over time may be driven by improvements in clinical trial reporting (CONSORT)
- passing familiarity may or may not allow full comprehension of results and implications by a hypothetical reader
- readers without familiarity with a particular method used in an article may be capable of understanding the overall results

(日) (同) (三) (三)

Implications for medical education Implications for mathematics and statistics educators GAISE



- importance of interpretation of conditional parameters from multiple regression
- comprehending generalizations of regression
- subtleties of interpretation of more complicated methods, particularly for observational studies

- 4 同 2 4 日 2 4 日 2

Implications for medical education Implications for mathematics and statistics educators GAISE

Implications

- Emerson and Colditz stated in 1992 that: an acquaintance with a few basic statistical techniques cannot give full statistical access to
 - research appearing in the Journal
- we believe that that this is even more true in 2006
- this increasing sophistication has potential implications for medical and statistical educators

・ 同 ト ・ ヨ ト ・ ヨ

Implications for medical education Implications for mathematics and statistics educators GAISE

What do physicians actually read?

• Example: Lange et al Folate, comprehension of multiple regression results

- 4 同 ト 4 ヨ ト 4 ヨ ト

Implications for medical education Implications for mathematics and statistics educators GAISE

What do physicians actually read?

- Example: Lange et al Folate, comprehension of multiple regression results
- Example: Vioxx (pain reliever) and side effects (cardiovascular) paper by Bombardier and colleagues

・ 戸 ト ・ ヨ ト ・ ヨ

Implications for medical education Implications for mathematics and statistics educators GAISE

Bombardier et al (NEJM 2000)

What if they only read the title and summary of the abstract?

Title: Comparison of Upper Gastrointestinal Toxicity of Rofecoxib (Vioxx) and Naproxen in Patients with Rheumatoid Arthritis

・ 戸 ト ・ ヨ ト ・ ヨ

Implications for medical education Implications for mathematics and statistics educators GAISE

Bombardier et al (NEJM 2000)

What if they only read the title and summary of the abstract?

Title: Comparison of Upper Gastrointestinal Toxicity of Rofecoxib (Vioxx) and Naproxen in Patients with Rheumatoid Arthritis

Conclusions: In patients with rheumatoid arthritis, treatment with rofecoxib, a selective inhibitor of cyclooxygenase-2, is associated with significantly fewer clinically important upper gastrointestinal events than treatment with naproxen, a nonselective inhibitor.

- 4 回 ト 4 ヨト 4 ヨト

Implications for medical education Implications for mathematics and statistics educators GAISE

What statistical training do physicians receive?

- none required for entry into medical school
- no requirement for course during medical school
- vague requirement for training in *accurate observation of biomedical phenomena and critical analyses of data*
- lots required for those pursuing research tracks
- is the base level of training for physicians sufficient?

・ 山 ト ・ ヨ ト ・

Implications for medical education Implications for mathematics and statistics educators GAISE



simple statistical methods concern us far more closely than many of the things that we are forced to learn in the six long years of the medical curriculum

(日) (同) (三) (三)

Implications for medical education Implications for mathematics and statistics educators GAISE

Windish et al, JAMA (2007)

Most residents in this study lacked the knowledge in biostatistics needed to interpret many of the results in published clinical research. Residency programs should include more effective biostatistics training in their curricula to successfully prepare residents for this important lifelong learning skill.

(人間) トイヨト イヨト

Implications for medical education Implications for mathematics and statistics educators GAISE

Recommendations for medical educators

- add a requirement of a statistics course prior to matriculation
- bolster required training in medical school (typically 6 weeks at present)
- teach in context (how to read and critique a journal paper)
- intensive short courses (Ambrosius and Manatunga, SIM 2002)
- focus on continuing education, use of journals to solidify and update knowledge

- 4 周 ト 4 戸 ト 4 戸 ト

Implications for medical education Implications for mathematics and statistics educators GAISE

Role of introductory and intermediate statistics in college

- Cobb and Moore reform efforts
- "more data, less lecture"
- "more poetry, less plumbing"
- key role of multiple regression (especially in intro)

Implications for medical education Implications for mathematics and statistics educators GAISE

Role of statistics, probability and data analysis in K-12

- needed to provide quantitative literacy to all citizens (not just doctors!)
- one of six strands in the National Council of Teachers of Mathematics (NCTM) standards from 1989 and 2000
- dramatic growth of AP statistics (100,000 students took the test last year)
- helps reinforce math skills (in context)
- NCTM new Focal Points for K-8 omit most of statistics, probability and data analysis

・ 同 ト ・ ヨ ト ・ ヨ

Introduction and motivation Increasing sophistication Implications and next steps Closing thoughts and discussion

Implications for mathematics and statistics educators

GAISE



Nicholas J. Horton

・ロッ ・ 一 ・ ・ ・ ・ The increasing sophistication of statistics in the NEJM

Implications for medical education Implications for mathematics and statistics educators GAISE

GAISE as a model

- GAISE (Guidelines for Assessment and Instruction in Statistics Education)
- framework for PreK 12 as well as college level statistics courses
- build overall foundation for additional sophistication
- allow physicians (as well as all citizens!) to assess uncertainty and variability, despite increasing complexity of statistical methods
- http://www.amstat.org/education/gaise

(日) (同) (三) (三)

Review References Discussion

Review

- reviewed findings from a survey of statistical methods used in articles in a medical journal widely read by clinicians
- described an increasing sophistication of statistical methods and quantitative reasoning
- considered examples where knowledge of probability and data analysis is needed to make informed decisions
- detailed the implications that increased use of sophisticated statistical methods have for quantitative skills education and for preparing citizens to be quantitatively literate.

(日) (同) (三) (三)

Review References Discussion

References

- Altman DG and Bland JM, *Improving doctors' understanding* of statistics, JRSS-A, (1991), 154:223-67
- Ambrosius WT and Manatunga AK, Intensive short courses in biostatistics for fellows and physicians, SIM (2002), 21:2739-2756
- Franklin C et al, *Guidelines for assessment and instruction in statistics education (GAISE) report*, (2007), American Statistical Association
- Horton NJ and Switzer SS, *Statistical methods in the Journal*, NEJM (2005), 353(18):1977-9
- Moore D and Cobb G, *Mathematics, statistics and teaching,* American Mathematical Monthly (1997), 104: 801-823

・ロト ・得ト ・ヨト ・ヨト

Review References Discussion

References

- NCTM Yearbook (2006), *Thinking and reasoning with data and chance*, National Council of Teachers of Mathematics
- Switzer SS and Horton NJ, What your physician should know about statistics (but perhaps doesn't), CHANCE (2007)
- West CP and Ficalora RD, Clinician attitudes towards biostatistics, Mayo Clinic Proceedings (2007), 82(8):939-943
- Windish DM, Huot SJ and Green ML, *Medicine residents'* understanding of the biostatistics and results in the medical literature, JAMA (2007), 298(9):1010-1022

Review References Discussion

Questions and discussion

Nicholas J. Horton The increasing sophistication of statistics in the NEJM

<ロ> <同> <同> < 同> < 同>

Review References Discussion

The increasing sophistication of statistics in the New England Journal of Medicine

Nicholas J. Horton

Smith College and University of Auckland

April 17, 2008, ASA TSHS webinar

nhorton@email.smith.edu http://www.math.smith.edu/~nhorton

- 4 同 6 4 日 6 4 日 6