How to Choose the Right Statistical Test for the Occasion

Lisa M Sullivan
Professor and Chair,
Department of Biostatistics BUSPH
July 7, 2010

Questions Before You Begin…

- What is your primary research question?
- Are you interested primarily in a relationship between an outcome and a risk factor or exposure?
- Are you interested in prediction of an outcome, using one or more risk factors or exposures?

Questions Before You Begin…

- What is the study design? (Cohort, Case/Control, RCT)
- What kinds of inferences can be made – what are the limitations?
- What are the outcome measures? Is there a primary outcome measure?
- What are the risk factors and exposures?

Questions Before You Begin…

- What is the nature of the outcome measure – Continuous, Categorical, Dichotomous or Time to Event?
- What is the primary effect measure?
- Are data correlated?

Example

- Is BMI a significant risk factor for spontaneous preterm delivery?

Analysis

- What is most appropriate study design?
- What is the nature of the outcome measure?
- What is the nature of the primary risk factor?
- What is effect measure?
- What is analysis method?

Analysis

- Conduct logistic regression analysis relating BMI to spontaneous preterm delivery

  $OR_{BMI} = 1.06$

  (95% CI: 1.04-1.08)

  p<0.0001
- Are you happy with this?

Interpretation…

Overweight BMI=25.0-29.9

Obese BMI $\geq$ 30

$OR_{Overwght} = 1.90$ (1.56-2.30), p<0.0001

$OR_{Obese} = 2.23$ (1.73-2.87), p<0.0001

- Done?

Suggestion: Explore Your Data

- Understand analytic sample - population at risk and outcome
- Generate descriptive statistics on outcome and risk factor
- Are there confounding factors?

Other Risk Factors/Confounders

- Prior preterm birth
- Maternal age
- Smoking
- Race
- Infection
- Alcohol & tobacco use
- Nutritional status

Confounding

- A distortion of the effect of the risk factor on outcome due to other factors
  - Confounder may account for part or all of observed effect, may mask effect
- How do we examine confounding?
  - Evaluate association of confounder with outcome
  - Evaluate association of confounder with primary risk factor of interest

Ways to Handle Confounding

- Design
  - Randomization
  - Matching
- Analysis
  - Stratification
  - Multivariable analysis/Statistical adjustment

Multivariable Models: BMI

	OR (95% CI)	p
Unadjusted	1.06 (1.04-1.08)	0.0001
Adj for Maternal Age	1.04 (1.02-1.06)	0.0001
Multivariable Adj*	1.02 (0.99-1.04)	0.1525

*Adjusted for prior preterm birth, maternal age, smoking and infection

Multivariable Models: Overweight/Obesity

	Overweight	Obese
Unadjusted	1.90 (1.56-2.30)	2.23 (1.73-2.87)
Adj for Mat Age	1.24 (1.02-1.52)	1.72 (1.33-2.33)
Multivariable Adj*	1.11 (0.91-1.36)	1.38 (1.04-1.78)

*Adjusted for prior preterm birth, maternal age, smoking and infection

Confounding

- Compare crude (unadjusted) measure of association with adjusted measure of association
  - If comparable, then no confounding
- Is confounding an issue here?
  - If so, want to explore which risk factor(s)?

What Tests to Use When

- Outcome Variable
  - Continuous (means) – birth weight
  - Discrete (proportions) – preterm labor
  - Time to Event (survival) – infant death
- Number of Groups
  - One
  - Two
  - > Two
- Independent or Dependent/Matched Groups

What Tests to Use When

- Continuous Outcome
  - 1 Group – CI, t Test for Mean
    - Historical control
  - 2 Independent Groups – CI, t Test for Difference in Means
  - 2 Dependent Groups – CI, t Test for Mean Difference (Post-Pre)
    - Focus on difference scores

What Tests to Use When

- Continuous Outcome
  - > 2 Independent Groups – ANOVA
    - Test for difference in means
    - Specific contrasts (2 at a time) but control for Type I error rate with multiple testing
  - > 2 Dependent Groups – Repeated Measures ANOVA
    - Repeated assessments over time
  - Multiple risk factors or exposures
    - Multivariable linear regression analysis

What Tests to Use When

- Dichotomous Outcome
  - 1 Group – CI, Z Test for Proportion
  - 2 Independent Groups – CI, Z Test for Difference in proportions
  - 2 Dependent Groups – McNemar's test for differences in proportions
  - > 2 Independent Groups – Chi-Square Test
  - Multiple risk factors or exposures
    - Multivariable logistic regression analysis

What Tests to Use When

- Time to Event
  - 1 Group - Kaplan Meier Estimate of Survival
  - 2+ Independent Groups – Log Rank Test for Differences in Survival
  - Multiple risk factors or exposures
    - Cox proportional hazards regression analysis

Common Mistakes

- Inefficient Design
  - A badly designed study can never be retrieved, a poorly analyzed study can usually be re-analyzed!
- Analytic Planning Issues
- Interpretation Issues

Which Design is Best

- Depends on the study question
- What is current knowledge on topic
- How common is disease (and risk factors)
- How long would study take, what are costs
- Ethical issues

Common Mistakes (cont'd)

- Misclassification of Outcome
  - Continuous (means)
  - Discrete (proportions)
    - Ordered categories, unordered categories, dichotomous (success/failure)
  - Time to event (survival time)

Common Mistakes (cont'd)

- Unit of analysis
  - Observations are repeated on the same unit but treated as independent
  - Observations are clustered, need to take into account structure in data

Common Mistakes (cont'd)

- Missing data
  - Suppose required number are enrolled but 20% drop out over the course of follow-up; What if 40% of the treatment group drop out and 0% of control drop out?
  - Patterns of missing data
  - Do everything possible to avoid missing data!

Common Mistakes (cont'd)

- Multiple testing
  - Each test has an associated Type I error (error rate per comparison, e.g. 5%)
  - Familywise error rate (likelihood of a false positive result over all comparisons)
  - Multiple comparisons procedures control familywise Type I error rate (e.g., Tukey, Dunnett)
  - Bonferroni correction

Common Mistakes (cont'd)

- Correlation Vs Cause and Effect

  Design

  Observational studies – correlation

  Experimental studies – cause and effect

  Timing

  Does A cause B or vice versa

Common Mistakes (cont'd)

- Lack of significance
  - Failure to show statistical significance is not equivalence (non-inferiority)
  - Must provide evidence of power when study fails to show statistical significance (equality or study is too small?)
- Determine sample size required BEFORE study launch

Common Mistakes (cont'd)

- Generalizabilty
  - Target population
  - Draw sample, analyze sample, make inferences back to target population

Magnitude of Effect

- Statistical significance ($p<0.05$) is only <u>one</u> way to interpret results
- Always look at magnitude of effect
- Consistency of effect in other studies
- Biologically plausible effect
- Dose-response relationship

Summary

- ☐ Determine appropriate study design
- ☐ Identify the types of variables you are evaluating
- ☐ Plan the appropriate analyses
  - Explore data
  - Run primary analysis
  - Assess consistency, plausibility

Study Variables

- ☐ Outcome – continuous, dichotomous, discrete, time to event
- ☐ Number of comparison groups
- ☐ Dependencies in the data

Summary

- ☐ Generate descriptive statistics for all variables, especially outcome and primary risk factor
- ☐ Obtain crude measures of association
- ☐ Perform stratified and adjusted analyses
- ☐ Does final result make sense, given all of the above and what you know from other studies?
- ☐ What are the limitations of analysis/ inferences?

Epidemiology

Soft Drink Consumption and Risk of Developing Cardiometabolic Risk Factors and the Metabolic Syndrome in Middle-Aged Adults in the Community

Ravi Dhingra, MD; Lisa Sullivan, PhD; Paul F. Jacques, PhD; Thomas J. Wang, MD; Caroline S. Fox, MD; James B. Meigs, MD, MPH; Ralph B. D'Agostino, PhD; J. Michael Gaziano, MD, MPH; Ramachandran S. Vasan, MD

Background—Consumption of soft drinks has been linked to obesity in children and adolescents, but it is unclear whether it increases metabolic risk in middle-aged individuals.

Methods and Results—We related the incidence of metabolic syndrome and its components to soft drink consumption in participants in the Framingham Heart Study (6039 person-observations, 3470 in women; mean age 52.9 years) who were free of baseline metabolic syndrome. Metabolic syndrome was defined as the presence of ≥3 of the following: waist circumference ≥35 inches (women) or ≥40 inches (men); fasting blood glucose ≥100 mg/dL; serum triglycerides ≥150 mg/dL; blood pressure ≥135/85 mm Hg; and high-density lipoprotein cholesterol <40 mg/dL (men) or <50 mg/dL (women). Multivariable models included adjustments for age, sex, physical activity, smoking, dietary intake of saturated fat, trans fat, fiber, magnesium, total calories, and glycemic index. Cross-sectionally, individuals consuming ≥1 soft drink per day had a higher prevalence of metabolic syndrome (odds ratio [OR], 1.48; 95% CI, 1.30 to 1.69) than those consuming <1 drink per day. On follow-up (mean of 4 years), new-onset metabolic syndrome developed in 765 (18.7%) of 4095 participants consuming <1 drink per day and in 474 (22.6%) of 2059 persons consuming ≥1 soft drink per day. Consumption of ≥1 soft drink per day was associated with increased odds of developing metabolic syndrome (OR, 1.44; 95% CI, 1.20 to 1.74), obesity (OR, 1.31; 95% CI, 1.02 to 1.68), increased waist circumference (OR, 1.30; 95% CI, 1.09 to 1.56), impaired fasting glucose (OR, 1.25; 95% CI, 1.05 to 1.48), higher blood pressure (OR, 1.18; 95% CI, 0.96 to 1.44), hypertriglyceridemia (OR, 1.25; 95% CI, 1.04 to 1.51), and low high-density lipoprotein cholesterol (OR, 1.32; 95% CI 1.06 to 1.64).

Conclusions—In middle-aged adults, soft drink consumption is associated with a higher prevalence and incidence of multiple metabolic risk factors. (***Circulation***. **2007;116:480–488.**)

Key Words: diabetes mellitus ■ metabolic syndrome ■ epidemiology ■ obesity ■ risk factors ■ carbonated beverages

TABLE 1. Baseline Characteristics of Participants According to Soft Drink Consumption (n=8997)

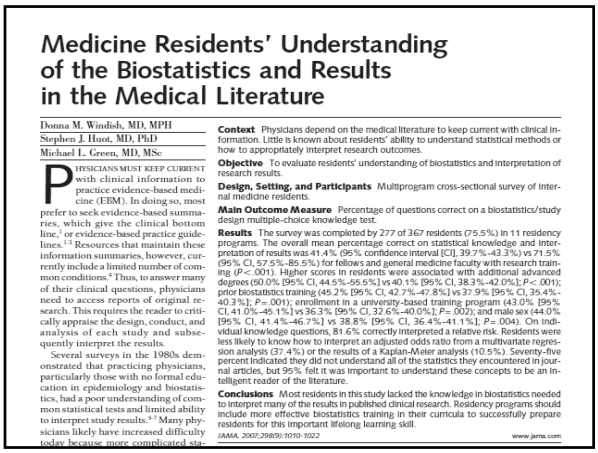
Characteristic	No. of Soft Drinks Consumed Per Day <1 (n=5840)	1 (n=1918)	≥2 (n=1239)	P*
Age, y	56±10	53±10	51±9	...
Men, %	42.8	50.2	53.4	...
Systolic BP, mm Hg	127±19	125±17	126±18	<0.0001
Diastolic BP, mm Hg	76±10	77±10	78±11	<0.0001
BP ≥130/85 mm Hg or on treatment, %	48.9	46.7	48.4	<0.0001
Hypertension, %	22.5	18.7	21.6	0.0014
Treatment for hypertension, %	18.9	16.1	17.6	0.0011
BMI, kg/m²	26.8±4.8	27.8±5.1	28.5±5.4	<0.0001
BMI ≥30 kg/m², %	20.9	27.1	32.1	<0.0001
Weight, kg	75.5±16.1	79.4±16.9	82.1±18.1	<0.0001
Waist circumference, in	36.0±5.6	36.9±5.7	37.8±6.1	<0.0001
Increased waist circumference, %†	33.9	37.2	41.1	<0.0001
Men	36.3	40.9	48.1	<0.0001¶
Women	32.0	33.4	33.2	<0.0001¶
Total cholesterol, mg/dL	206±37	204±37	202±38	0.72
Low-density lipoprotein cholesterol, mg/dL	129±34	128±33	127±34	0.30
Triglycerides, mg/dL	127±83	141±119	148±118	<0.0001
High triglycerides, %‡	28.3	32.7	35.9	<0.0001
HDL-C, mg/dL	52±16	50±15	47±14	<0.0001
Low HDL-C, %§	34.8	38.7	46.1	<0.0001
Men	37.5	42.0	45.1	<0.0001¶
Women	32.8	35.5	47.2	<0.0001¶
Blood sugar, mg/dL	97±21	99±26	105±39	<0.0001

TABLE 2. Cross-Sectional Relationships of Soft Drink Consumption With Prevalence of Metabolic Syndrome

Soft Drink Consumption, Servings/d	Metabolic Syndrome, n	No. at Risk*	Age- and Sex-Adjusted OR (95% CI)	Multivariable Adjusted OR (95% CI)†
Model I: any soft drink (regular or diet); data from all 3 examinations (4, 5, and 6; n=8997)				
None	1697	5840	Referent	Referent
1	618	1918	1.18 (1.06 to 1.33)	1.38 (1.19 to 1.61)
≥2	462	1239	1.43 (1.24 to 1.66)	1.67 (1.38 to 2.01)

TABLE 3. Multiple Logistic Regression Examining Soft Drink Consumption and Incidence of Metabolic Syndrome (n=6154)

Soft Drink Consumption, Servings/d	Metabolic Syndrome, n	No. at Risk*	Age- and Sex-Adjusted OR (95% CI)	Multivariable-Adjusted OR (95% CI)†
Model I: any soft drink (regular or diet); data from all 3 examinations (4, 5, and 6; n=6154)				
None	717	4033	Referent	Referent
1	267	1259	1.34 (1.14 to 1.58)	1.53 (1.24 to 1.89)
≥2	166	747	1.46 (1.20 to 1.78)	1.29 (0.98 to 1.70)



Medicine Residents' Understanding of the Biostatistics and Results in the Medical Literature

Donna M. Windish, MD, MPH
Stephen J. Huot, MD, PhD
Michael L. Green, MD, MSc

PHYSICIANS MUST KEEP CURRENT with clinical information to practice evidence-based medicine (EBM). In doing so, most prefer to seek evidence-based summaries, which give the clinical bottom line,[1] or evidence-based practice guidelines.[1,2] Resources that maintain these information summaries, however, currently include a limited number of common conditions.[3] Thus, to answer many of their clinical questions, physicians need to access reports of original research. This requires the reader to critically appraise the design, conduct, and analysis of each study and subsequently interpret the results.

Several surveys in the 1980s demonstrated that practicing physicians, particularly those with no formal education in epidemiology and biostatistics, had a poor understanding of common statistical tests and limited ability to interpret study results.[4-7] Many physicians likely have increased difficulty today because more complicated sta-

Context Physicians depend on the medical literature to keep current with clinical information. Little is known about residents' ability to understand statistical methods or how to appropriately interpret research outcomes.

Objective To evaluate residents' understanding of biostatistics and interpretation of research results.

Design, Setting, and Participants Multiprogram cross-sectional survey of internal medicine residents.

Main Outcome Measure Percentage of questions correct on a biostatistics/study design multiple-choice knowledge test.

Results The survey was completed by 277 of 367 residents (75.5%) in 11 residency programs. The overall mean percentage correct on statistical knowledge and interpretation of results was 41.4% (95% confidence interval [CI], 39.7%-43.3%) vs 71.5% (95% CI, 57.5%-85.5%) for fellows and general medicine faculty with research training ($P < .001$). Higher scores in residents were associated with additional advanced degrees (50.0% [95% CI, 44.5%-55.5%] vs 40.1% [95% CI, 38.3%-42.0%]; $P < .001$); prior biostatistics training (45.2% [95% CI, 42.7%-47.8%] vs 37.9% [95% CI, 35.4%-40.3%]; $P = .001$); enrollment in a university-based training program (43.0% [95% CI, 41.0%-45.1%] vs 36.3% [95% CI, 32.6%-40.0%]; $P = .002$); and male sex (44.0% [95% CI, 41.4%-46.7%] vs 38.8% [95% CI, 36.4%-41.1%]; $P = .004$). On individual knowledge questions, 81.6% correctly interpreted a relative risk. Residents were less likely to know how to interpret an adjusted odds ratio from a multivariate regression analysis (37.4%) or the results of a Kaplan-Meier analysis (10.5%). Seventy-five percent indicated they did not understand all of the statistics they encountered in journal articles, but 95% felt it was important to understand these concepts to be an intelligent reader of the literature.

Conclusions 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.

JAMA. 2007;298(9):1010-1022
www.jama.com