The Effect on Benefit Estimates of Discarding Data

Richard B. Belzer
Good Intentions Paving Co.
luke@goodintentionspaving.co
R. Jeffrey Lewis
ExxonMobil Biomedical Sciences
r.jeffrey.lewis@exxonmobil.com

Society for Benefit Cost Analysis
2016 Annual Conference
March 17-18, 2016
George Washington University

Presentation Outline

- Normal spirometric function
- How spirometry data are usually collected
- An exploratory data analysis (EDA)
- Simulation expanding on the EDA
- Application to ozone
  - Chamber studies
  - Observational epidemiology

NORMAL SPIROMETRIC FUNCTION

Normal FEV_1 in Never-Smoking Asymptomatic Adults

- FEV_1 is log-normally distributed function of sex, age and height
  - Height (H) effect is linear in logs
  - Age (A) effect is non-linear in logs
  - Separate eqs. estimated by sex, age ≤ 25, age > 25
- Ref. equation for normal 59-year old 5'8" M
  - $e^{(-8.240 + 1.9095 \ln(H\ [\text{in cm}]) - 0.0037\ A - 0.000033\ A^2)} = 3.55$

Source: Brändli O, Schindler C, Künzli N, Keller R, Perruchoud A. 1996. Lung function in healthy never smoking adults: reference values and lower limits of normal of a Swiss population. Thorax 51:277-283.

HOW SPIROMETRY DATA ARE TYPICALLY COLLECTED

ATS Spirometry Protocol

1. Conduct maneuvers (1), (2) and (3).
2. If FEV_1 and FVC are within 150 mL for any pair, quit and record maximum. If else, continue.
3. Conduct an additional maneuver.
4. If FEV_1 and FVC are within 150 mL for any pair, quit and record maximum. If else, return to step 3.
5. If the number of maneuvers performed equals eight, quit and discard subject.

- Note: the ATS protocol does not explain why the maximum value is used as the representative indicator of pulmonary function.

ATS Recommended Clinical Interpretation of Spirometry Data

Clinical Interpretation	Criterion
"May be a physiological variant" [a]	≥ 100% of predicted
"Mild" [b]	70-100% of predicted
"Moderate" [b]	60-69% of predicted
"Moderately Severe" [b]	50-59% of predicted
"Severe" [b]	35-49% of predicted
"Very Severe" [b]	< 35% of predicted

Sources:
(a) American Thoracic Society. 1991. Lung Function Testing: Selection of Reference Values and Interpretative Strategies. Am Rev Resp Dis 144:1202-1218.
(b) Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, et al. 2005. Interpretative Strategies for Lung Function Tests. Eur Res J 26:948-968.

ATS Definition of 'Significant Change' in Pulmonary Function

Differences by Time Period	FVC	FEV_1
Within day (normal)	≥ 5%	≥ 5%
Within day (COPD)	≥ 11%	≥ 13%
Week to week (normal)	≥ 11%	≥ 12%
Week to week (COPD)	≥ 20%	≥ 20%
Year to year	≥ 15%	≥ 15%

Source: Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, et al. 2005. Interpretative strategies for lung function tests. European Respiratory Journal 26:948-968, Table 12

AN EXPLORATORY DATA ANALYSIS

Exploratory Data Analyis

- 1 subject (RBB)
- 15 FEV_1 tests performed over 12 days under identical conditions (except time of day) following ATS Protocol
- 8 maneuvers per test, 1-2 minutes apart
- All data recorded

Repeatability Censoring Prevents Collection of Valid Data

Test 6

Maneuver	Max	Max	Max	Max	Max	Max	Max
1							
2	2.43	2.43					
3							
4			2.54	2.54	2.54	2.54	
5							
6							
7							
8							2.57

Test 15

Maneuver	Max	Max	Max	Max	Max	Max	Max
1	2.37						
2							
3		2.49	2.49	2.49			
4							
5							
6					2.54		
7						2.57	2.57
8							

Repeatability Censoring Prevents Collection of Valid Data

Test 8

Maneuver	L/sec	Max M=3	Max M=4	Max M=5	Max M=6	Max M=7	Max M=8
1	2.78	2.78					
2	2.75						
3	2.77						
4	2.78						
5	2.89			2.89			
6	3.00				3.00	3.00	
7	2.91						
8	3.20						3.20

Repeatability Censoring Discards Valid Data

Test #2 With Censor

Maneuver	L/sec	Max M=3	Max M=4	Max M=5	Max M=6	Max M=7	Max M=8
1	2.09	2.20					
2	2.20						
3	2.51						
4	2.29		2.29	2.29	2.29	2.29	2.29
5	2.17						
6	2.15						
7	2.09						
8	2.17						

Test #2 Without Censor

Maneuver	L/sec	Max M=3	Max M=4	Max M=5	Max M=6	Max M=7	Max M=8
1	2.09	2.20					
2	2.20						
3	2.51	2.51	2.51	2.51	2.51	2.51	2.51
4	2.29						
5	2.17						
6	2.15						
7	2.09						
8	2.17						

Potential Effects on Clinical Classification

	Min Maneuvers 100 mL Censor	Min Maneuvers 150 mL Censor	Min Maneuvers 200 mL Censor	Min Maneuvers No Censor
Potential Misclassifications	--	3/15	0/15	1/15
Classification Changes	--	1/3	--	1/1

	8 Maneuvers 100 mL Censor	8 Maneuvers 150 mL Censor	8 Maneuvers 200 mL Censor	8 Maneuvers No Censor
Potential Misclassifications	--	0/15	2/15	0/15
Classification Changes	--	0/0	0/2	0/0

HOW DO ATS PROTOCOL AND FULL DATA SETS COMPARE?

Inter-test Variability from Literature

Study	Subjects	Tests/ Subject	Techs/ Subject	Retest period days	Mean Age yrs	CV^t FVC %	CV^t FEV_1 %
SAPALDIA							
8 teams	13	8	8	2	24	2.7	3.3
8 devices	13	8	1	1	24	2/0	2.2
McCarthy et al	12	10	1	1	27	2.5	2.5
Cochrane et al	9	10	1	1	24	1.8	2.3
Nickerson et al	15	5-12	1	1	18	3.5	3.6
Lebowitz et al	10	60	1	25-35	34	3.5	3.6
Rozas & Goldman	15	5	1	5	?	2.8	2.8
Groth et al	112	2	1-2	15-180	47	4.9	4.7

Sources: (b) Künzli N, Ackermann-Liebrich U, Keller R, Perruchoud AP, Schindler C, Team S. 1995. Variability of FVC and FEV1 due to technician, team, device and subject in an eight centre study. European Respiratory Journal 8:371–376, Table 3.

Inter- and Intra-test Variability in EDA

Statistic	ATS Protocol 100 mL Censor (L/s)	ATS Protocol 150 mL Censor [L/s]	ATS Protocol 200 mL Censor [L/s]	8 Maneuvers No Censor [L/s]
Highest Max	2.78	3.00	3.20	3.20
Average Max	2.53	2.60	2.55	2.64
Lowest Max	2.29	2.29	2.20	2.44
St Dev	0.153	0.161	0.161	0.196
Excluded Maxima	24/120 (20%)	11/120 (9.2%)	6/120 (5.0%)	0/120 (0.0%)
CV^t	6.1%	6.4%	6.3%	5.0%
CV_m	N/A	N/A	N/A	3.3% (1.2-6.2%)

Inter- and Intra-test Variability: EDA, Literature & NHANES 2009-10

	EDA (ATS 2005)	EDA (Unconstrained)	Literature [a]	NHANES 2009-10
Inter-test Variability (CV^t)	6.3%	5.0%	2.2 to 4.7%	N/A
Intra-test Variability (CV_m)	N/A	3.3% ± Range: 1.3-6.2%	N/A	6% (Sample)

(a) Kunzli et al 1995

Baseline Simulation Parameters

- 10k tests, 8 maneuvers/each
- Predicted max FEV_1 (from Brändli et al eq., normal, non-smoking 59-y 5'8" M)
- Inter-test coefficient of variation [CV^t] (from Künzli et al)
- Intra-test coefficient of variation [CV_m] (from NHANES 2009-10)

Note: All parameters user-adjustable.

Baseline Simulation Model

2 Stage Simulation		Model	Coefficient of Variation	StDev of Maxima
1	Across 10k tests *t*	X^t = NORMDIST (RAND(), 3.55, σ^t)	CV^t = 3%	σ^t = 0.11
2	Across 8 maneuvers *m* per test	X_m^t = NORMDIST (RAND(), X^t, σ_m)	CV_m = 6%	σ_m = 0.21

Interquartile Range of Artifactual Bias in ATS Protocol

Sensitivity Analysis

Repeatability Censor	*Intra-Test Coefficient of Variation (CV_m)*		
	3%	*6%*	*9%*
100 mL	130—110 mL/sec 3.7—3.0%	280—250 ml/sec 6.8—8.2%	450—400 ml/sec 13—11%
150 mL	90—110 ml/sec 2.9—2.3%	270—230 mL/sec 7.9—6.4%	440—370 mL/sec 13—10%
200 mL	90—80 mL/sec 2.0—2.6%	260—210 mL/sec 7.3—5.7%	420—350 mL/sec 9.6—12%

All biases reported with two significant digits (±10 ml/sec).
All percentages reported with two significant digits (±.05%).
Interquartile range: 25^{th}—75^{th} percentile of simulated distribution.

APPLICATION TO CHAMBER STUDY DATA

Schelegle et al (2009)
Authors' Abstract and Conclusions

- N=31, 16F + 15M, 18-25y, five 6.6-hour chamber exposures
- Results at the end of each protocol:

Ozone Exposure (ppb)	Mean + SD FEV_1 Decrement (%)
0	+0.80 ± 0.90%
60	–2.72 ± 1.48%
70	–5.34 ± 1.42%
80	–7.02 ± 1.60%
87	–11.42 ± 2.20%

- Conclusions: Inhalation of 70 ppb O_3 for 6.6 hours is sufficient to induce statistically significant decrements in FEV_1 in healthy young adults.

Schelegle et al (2009)

Are Results at Mean within Artifactual Bias?

Females

O_3 (ppb)	Mean FEV_1 Decline (%)	% Tests Where Mean FEV_1 Decline > Bias [a]		
		$CV_m = CV^t$	$CV_m = CV^t/2$	$CV_m = CV^t/10$
0	+0.80	27%	53%	79%
60	-2.72	29%	54%	79%
70	-5.34	32%	61%	81%
80	-7.02	38%	66%	82%
87	-11.42	52%	78%	84%
Zero bias		*3%*	*8%*	*15%*

FEV_1 mean (SD) = 4.16 (0.19); repeatability censor = 150 mL/sec; CV^t = 3%.

[a] Proportion of tests in which both draws have zero bias in parentheses.

Schelegle et al (2009)

Are Results at Mean within Artifactual Bias?

Males

O_3 (ppb)	Mean FEV_1 Decline (%)	% Tests Where Mean FEV_1 Decline > Bias [a]		
		$CV_m = CV^t$	$CV_m = CV^t/2$	$CV_m = CV^t/10$
0	+0.80	26%	52%	79%
60	-2.72	30%	57%	81%
70	-5.34	39%	66%	83%
80	-7.02	45%	73%	83%
87	-11.42	65%	87%	86%
Zero bias		*3%*	*8%*	*14%*

FEV_1 mean (SD) = 5.72 (0.19); repeatability censor = 150 mL/sec; CV^t = 3%.

[a] Proportion of tests in which both draws have zero bias in parentheses.

APPLICATION TO OBSERVATION EPIDEMIOLOGIC STUDIES

Use of ATS Protocol in Air Pollution Research

- Typical research objective: estimate differences and allocate proper fraction to competing causes.
- But, within-subject variability is ignored:
  - Assumes within-subject inter-test variability (CV^t) = 0
  - Assumes within-subject intra-test variability (CV_m) = 0
- Protocols
  - Do not collect data sufficient to estimate CV^t or CV_m
  - Discard valid data outside of repeatability censor
- Like chamber studies, differences described as statistically significant may be smaller than artifactual bias in the protocol

Studies Used for O_3 Benefits

Table 6-7. Health Endpoints and Epidemiological Studies Used to Quantify Ozone-Related Health Impacts [a]

Endpoint	Study	Study Population	Relative Risk or Effect Estimate (β) (with 95th Percentile Confidence Interval or SE)
Premature Mortality			
Premature mortality—short-term	*Smith et al. (2009)*	All ages	β = 0.00032 (0.00008)
	Zanobetti and Schwartz (2008)		β = 0.00051 (0.00012)
Premature respiratory mortality-long-term	*Jerrett et al. (2009)*	>29 years	β = 0.003971 (0.00133)
Hospital Admissions			
Respiratory	Pooled estimate: *Katsouyanni et al. (2009)*	> 65 years	β = 0.00064 (0.00040) penalized splines
Asthma-related emergency department visits	Pooled estimate: *Glad et al. (2012)*	0-99 years	β = 0.00306 (0.00117)
	Ito et al. (2007)		β = 0.00521 (0.00091)
	Mar and Koenig (2010)		β = 0.01044 (0.00436) (0-17 yr olds) β = 0.00770 (0.00284) (18-99 yr olds)
	Peel et al. (2005)		β = 0.00087 (0.00053)
	Sarnat et al. (2013)		β = 0.00111 (0.00028)
	Wilson et al. (2005)		RR = 1.022 (0.996 – 1.049) per 25
Other Health Endpoints			
Asthma exacerbation	Pooled estimate: [b] *Mortimer et al. (2002)*	6–18 years	β = 0.00929 (0.00387)
	Schildcrout et al. (2006)		β = 0.00222 (0.00282)
School loss days	Pooled estimate: Chen et al. (2000)	5-17 years	β = 0.015763 (0.004985)
	Gilliland et al. (2001)		β = 0.007824 (0.004445)
Acute respiratory symptoms (MRAD)	Ostro and Rothschild (1989)	18–65 years	β = 0.002596 (0.000776)

[a] Studies highlighted in *red* represent updates incorporated since the 2008 ozone NAAQS RIA (U.S. EPA, 2008a).
[b] The original study populations were 5 to 12 years for Schildcrout et al. (2006) and 5-9 years for the Mortimer et al. (2002) study. Based on advice from the SAB-HES, we extended the applied population to 6-18 years for all three studies, reflecting the common biological basis for the effect in children in the broader age group. See: U.S. EPA-SAB (2004a) and NRC (2002).

Role of FEV_1 in O_3 Benefits

Health Endpoint	Study	Is FEV_1 Decrement a Key Event in Mode of Action?
MORTALITY	Smith et al (2009) – Short term	No MOA given (non-accidental, all-cause)
	Zanobetti & Schwartz (2008) – Short term	Yes (by inference)
	Jerrett et al (2009) – Long term	Yes (respiratory) Yes but (protective cardiopulmonary)
HOSPITAL VISTIS OR ADMISSIONS	Peel et al (2005)	Yes (asthma, COPD, URI, pneumonia)
	Wilson et al (2005)	Yes (asthma, other respiratory)
	Katsouyanni et al (2009)	Yes (respiratory)
	Glad et al (2007)	Yes (asthma)
	Ito et al (2007)	Yes (asthma)
	Mar & Koenig (2010)	Yes (asthma)
	Sarnat et al (2013)	Yes (asthma)
ASTHMA EXACERBATION	Ostro & Rothschild (1989)	Yes
	Mortimer et al (2002)	Yes
	Schildcrout et al (2006)	Yes

Studies Used for $PM_{2.5}$ Co-Benefits

Table 6-8. Health Endpoints and Epidemiological Studies Used to Quantify $PM_{2.5}$-Related Health Impacts [a]

Endpoint	Study	Study Population	Relative Risk or Effect Estimate (β) (with 95th Percentile Confidence Interval or SE)
Premature Mortality			
Premature mortality—cohort study, all-cause	*Krewski et al. (2009)*	> 29 years	RR = 1.06 (1.04–1.06) per 10 μg/m³
	Lepeule et al. (2012)	> 24 years	RR = 1.14 (1.07–1.22) per 10 μg/m³
Premature mortality—all-cause	Woodruff et al. (1997)	Infant (< 1 year)	OR = 1.04 (1.02–1.07) per 10 μg/m³
Chronic Illness			
Nonfatal heart attacks	Peters et al. (2001)	Adults (> 18 years)	OR = 1.62 (1.13–2.34) per 20 μg/m³
	Pooled estimate:		
	Pope et al. (2006)		β = 0.00481 (0.00199)
	Sullivan et al. (2005)		β = 0.00198 (0.00224)
	Zanobetti et al. (2009)		β = 0.00225 (0.000591)
	Zanobetti and Schwartz (2006)		β = 0.0053 (0.00221)
Hospital Admissions			
Respiratory	*Zanobetti et al. (2009)*—ICD 460-519 (All respiratory)	> 64 years	β=0.00207 (0.00446)
	Kloog et al. (2012)—ICD 460-519 (All Respiratory		β=0.0007 (0.000961)
	Moolgavkar (2000)—ICD 490–496 (Chronic lung disease)	18–64 years	1.02 (1.01–1.03) per 36 μg/m³
	Babin et al. (2007)—ICD 493 (asthma)	< 19 years	β=0.002 (0.004337)
	Sheppard (2003)—ICD 493 (asthma)	< 18	RR = 1.04 (1.01–1.06) per 11.8 μg/m³
Cardiovascular	Pooled estimate:	> 64 years	
	Zanobetti et al. (2009)—ICD 390-459 (all cardiovascular)		β=0.00189 (0.000283)
	Peng et al. (2009)—ICD 426-427; 428; 430–438; 410–414; 429; 440–449 (Cardio-, cerebro- and peripheral vascular disease)		β=0.00068 (0.000214)
	Peng et al. (2008)—ICD 426-427; 428; 430–438; 410–414; 429; 440–449 (Cardio-, cerebro- and peripheral vascular disease)		β=0.00071 (0.00013)
	Bell et al. (2008)—ICD 426–427; 428; 430–438; 410–414; 429; 440–449 (Cardio-, cerebro- and peripheral vascular disease)		β=0.0008 (0.000107)
	Moolgavkar (2000)—ICD 390–429 (all cardiovascular)	20–64 years	RR=1.04 (t statistic: 4.1) per 10 μg/m³
Asthma-related emergency department visits	Pooled estimate:	All ages	
	Mar et al. (2010)		RR = 1.04 (1.01–1.07) per 7 μg/m³
	Slaughter et al. (2005)		RR = 1.03 (0.98–1.09) per 10 μg/m³
	Glad et al. (2012)		β=0.00392 (0.002843)
Other Health Endpoints			
Acute bronchitis	Dockery et al. (1996)	8–12 years	OR = 1.50 (0.91–2.47) per 14.9 μg/m³

Role of FEV_1 in $PM_{2.5}$ Co-Benefits

Health Endpoint	Study	Is FEV_1 Decrement a Key Event in Mode of Action?
MORTALITY (ALL CAUSES)	Krewski et al (2009)	No MOA given (adults > 29 years)
	Lepeule et al (2012)	No MOA given (adults > 24 years)
	Woodruff et al (1997)	No MOA given (infants < 1 year)
NONFATAL HEART ATTACKS	Peters et al (2001)	Yes
	Pope et al (2006)	Yes
	Sullivan et al (2005)	Yes
	Zanobetti et al (2009)	Yes
	Zanobetti and Schwartz (2006)	Yes
HOSPITAL ADMISSIONS	Zanobetti et al (2009)	Yes (all respiratory)
	Kloog et al (2012)	Yes (all respiratory)
	Moolgavkar (2000)	Yes (chronic lung disease)
	Babin et al (2007)	Yes (asthma)
	Sheppard (2003)	Yes (asthma)

Role of FEV_1 in $PM_{2.5}$ Co-Benefits Cont'd

Health Endpoint	Study	Is FEV_1 Decrement a Key Event in Mode of Action?
HOSPITAL VISITS OR ADMISSIONS	Zanobetti et al (2009)	Yes (asthma, COPD, URI, pneumonia)
	Peng et al (2009)	Yes (asthma, other respiratory)
	Peng et al (2008)	Yes (respiratory)
	Bell et al (2008)	Yes (asthma)
	Moolgavkar (2000)	Yes (asthma)
	Mar et al (2010)	Yes (asthma)
	Slaughter et al (2005)	Yes (asthma)
	Glad et al (2012)	Yes (asthma)
ACUTE BRONCHITIS	Dockery et al (1996)	Yes

CONCLUSIONS AND NEXT STEPS

Future Work

- Empirically estimate CV^t and CV_m in interesting subpopulations.
- Propose revision of protocol to account for CV^t and CV_m.
- Extend mean chamber study analysis to individual subjects and re-estimate.
- Re-estimate results for observational epi studies. (Depends on authors making data available.)
- Re-estimate benefits.

References

- Belzer RB and Lewis RJ. [ ]. Does the standard pulmonary function testing protocol produce biased estimates? An exploratory analysis and simulation study (in peer review).
- Brändli O, et al., 1996. Lung function in healthy never smoking adults: reference values and lower limits of normal of a Swiss population, 51 Thorax 277-283.
- Künzli N, et al., 1995. Variability of FVC and FEV_1 due to technician, team, device and subject in an eight centre study, 8 Eur Respir J 371-375.
- Miller MR, Hankinson J, Brusasco V, et al. 2005. Standardisation of spirometry. Eur Respir J, 26(2), 319-338.
- National Institute of Occupational Safety and Health (NIOSH). 1989. NIOSH spirometry training guide. NIOSH, Morgantown, WV.
- Pellegrino R, Viegi G, Brusasco V, et al. 2005. Interpretative strategies for lung function tests. Eur Respir J 26(2): 948-68.
- U.S. Environmental Protection Agency. 2015. Regulatory Impact Analysis of the final revisions to the National Ambient Air Quality Standard for ground-level ozone.

Questions?