Discarding Data Overstates Risk Estimates from Exposure to Ambient Air Pollutants

Richard B. Belzer
Good Intentions Paving Co.
rbbelzer@post.harvard.edu
R. Jeffrey Lewis
ExxonMobil Biomedical Sciences
r.jeffrey.lewis@exxonmobil.com

Presentation Outline

- Normal spirometric function
- How spirometry data are collected
- An exploratory data collection
- How do ATS protocol and full data sets compare?
- Conclusions and next steps

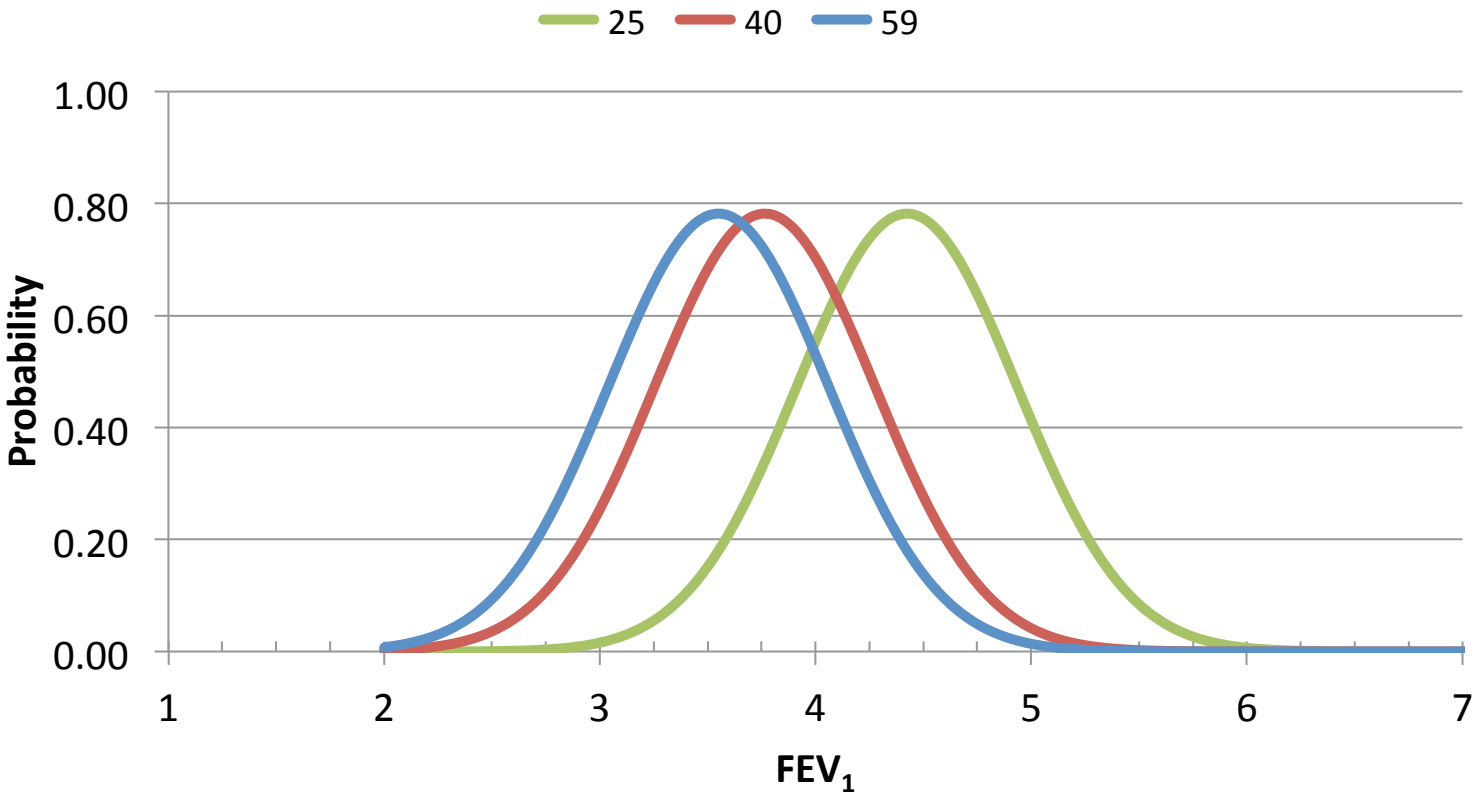
NORMAL SPIROMETRIC FUNCTION

Normal FEV_1 in Never-Smoking Adults without Respiratory Symptoms

- 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
- Reference equation for men age > 25
  - $e^{(-8.240 + 1.9095 \ln(H\ [\text{in cm}]) - 0.0037\ A - 0.000033\ A^2)} = 3.72$

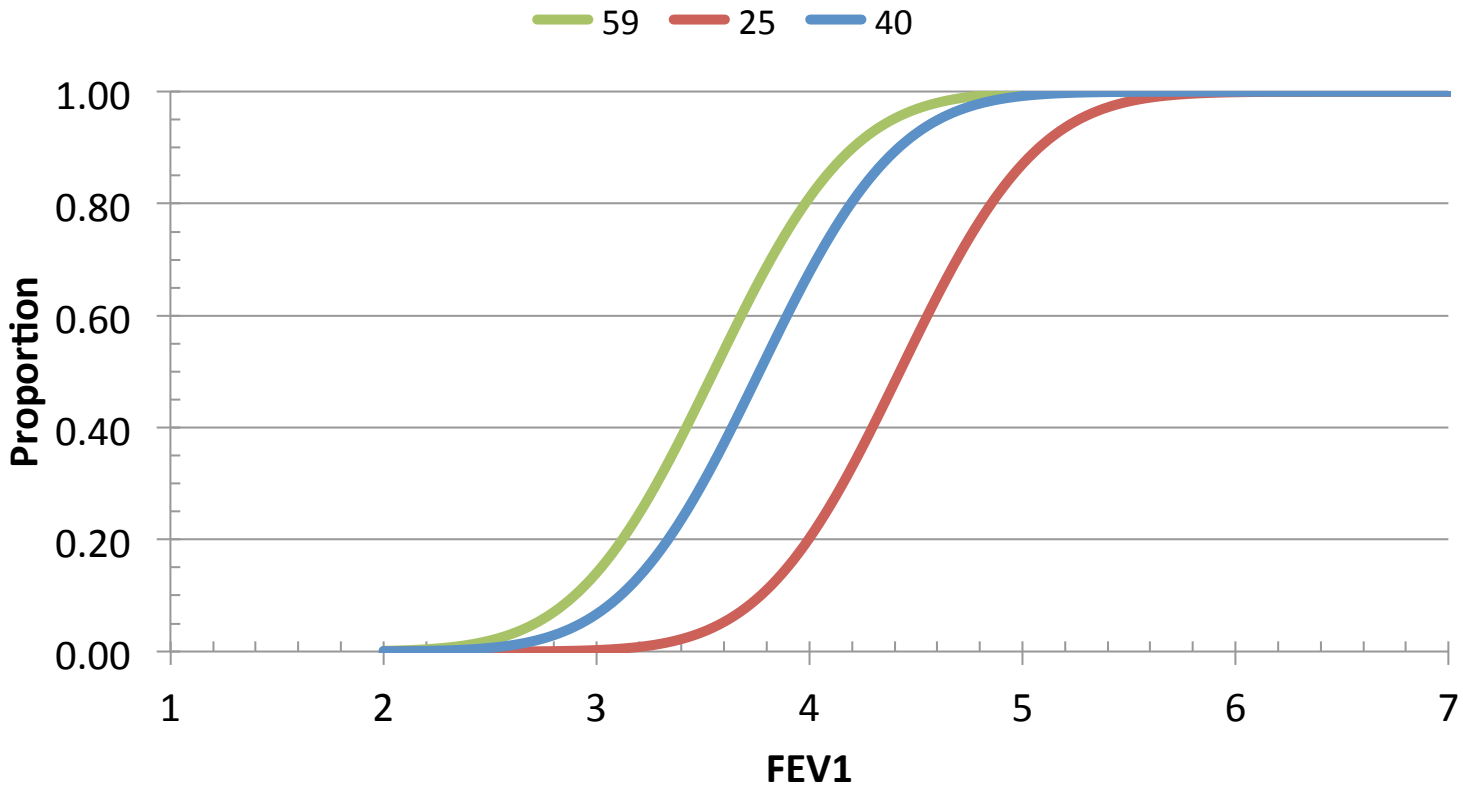
Source: Brändli O, Schindler C, Künzli N, Keller R, Perruchoud A. 1996. Lung function in healthy never smoking adults: reference values and lower limits of normal of a Swiss population. Thorax 51:277-283.

FEV_1 for Males Ages 25, 40 & 59



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.

FEV_1 CDF for Males Ages 25, 40 & 59



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 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.

Maneuvers and Censors in ATS Protocols, 1979-date

Year	Man-euvers	Censor	Notes
1979 1987 1991 1993	3 to 8	Max 5% of highest or 100mL	Earliest protocol located
1995	3 to 8	200 mL	Increased to 200 mL based on Hankinson & Bang (1991): Older "ATS reproducibility criterion, based on a percentage of the FVC and FEV_1, may inappropriately classify a higher percentage of subjects with smaller heights and lung volumes as having a nonreproducible test."
2005	3 to 8	150 mL	No explanation

ATS Recommended Clinical Interpretation of Spirometry Data

Clinical Interpretation	Criteria for Admission
"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	FEV1
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

EPA Definition of 'Adverse Effect'

Type of Difference	FVC	FEV_1
Across tests	?	Reversible decrement ≥ 10%
Population, subpopulation or sample	?	Subset of sample

Source: U.S. Environmental Protection Agency. 2015. National Ambient Air Quality Standards for Ozone; Final Rule. Fed. Reg. 80:65292-65468.

ATS Definition of 'Adverse Effect'

Year	Definition	Notes
1985	One or more of the following: 1. interference with normal activity 2. Episodic respiratory illness 3. Incapacitating illness 4. Permanent respiratory injury 5. Progressive respiratory dysfunction.	• 'The term "adverse health effect" is … is often used indiscriminately and loosely.' • 'As methodologic and analytic techniques become more sophisticated and sensitive, effects will be detected that may be clinically equivocal in significance.'
2014	• < 5^{th} percentile of reference value • 15% FEV_1 decline (plus age-related loss) ≥ 5 yrs	• 'Abnormal' • 'Excessive' FEV_1 decline; assumes 6% within-person variation across tests

AN EXPLORATORY DATA COLLECTION

Exploratory Data Collection

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

ATS Stopping Rule 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							

ATS 'Acceptability Criterion' Discards Valid Data

Test #2 With Censor

Maneuver	Max	Max	Max	Max	Max	Max	Max
1	2.20	2.20					
2							
3							
4			2.29	2.29	2.29	2.29	2.29
5							
6							
7							
8							

Test #2 Without Censor

Maneuver	Max	Max	Max	Max	Max	Max	Max
1	2.20	2.20					
2							
3		2.51	2.51	2.51	2.51	2.51	2.51
4			2.29	2.29	2.29	2.29	2.29
5							
6							
7							
8							

For 15 Tests, Maneuver with Recorded Value Differs under Different Stopping and Censoring Rules

Rules	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
ATS stop + 100 mL/sec	1	2	3	1	2	2	1	1	3	2	2	3	2	2	3
ATS stop +150 mL/sec	1	2	1	3	2	2	3	1	3	3	2	3	2	2	3
No stop No censor	4	3	1	3	2	8	3	8	5	6	8	3	2	2	7

ATS Stopping Rule and 150 ml Censor Affected 3 of 15 Clinical Classifications

Test	A. ATS Protocol; minimum maneuvers; 150 ml censor	ATS Classification	B. ATS Protocol; 8 maneuvers; 150 ml censor	ATS Classification	C. ATS Protocol; 8 maneuvers; No censor	ATS Classification	B - A	C - A
1	2.52	Moderate	2.57	Moderate	2.57	Moderate	0.05	0.05
2	2.20	Moderately Severe	2.29	Moderate	2.51	Moderate	0.09	0.31
3	2.69	Mild	2.69	Mild	2.69	Mild	0.00	0.00
4	2.55	Moderate	2.66	Mild	2.66	Mild	0.11	0.11
5	2.69	Mild	2.69	Mild	2.69	Mild	0.00	0.00
6	2.43	Moderate	2.57	Moderate	2.57	Moderate	0.14	0.14
7	2.46	Moderate	2.59	Mild	2.59	Mild	0.13	0.13
8	2.78	Mild	3.00	Mild	3.20	Mild	0.22	0.42
9	2.69	Mild	2.69	Mild	2.86	Mild	0.00	0.17
10	2.40	Moderate	2.46	Moderate	2.46	Moderate	0.06	0.06
11	2.43	Moderate	2.45	Moderate	2.45	Moderate	0.02	0.02
12	2.46	Moderate	2.54	Moderate	2.54	Moderate	0.08	0.08
13	2.44	Moderate	2.44	Moderate	2.44	Moderate	0.00	0.00
14	2.77	Mild	2.77	Mild	2.77	Mild	0.00	0.00
15	2.37	Moderate	2.57	Moderate	2.57	Moderate	0.20	0.20
Max	2.78		3.00		3.20		0.22	0.42
Avg	2.53		2.60		2.64		0.07	0.11
StDev	0.166		0.166		0.196		0.074	0.123
CoV	6.6%		6.4%		7.4%		101%	109%
% Δ							8%	15%

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-test and Inter-maneuver Variability in Exploratory Data

Statistic	ATS Protocol [L/sec]	ATS Protocol with 8 Maneuvers [L/sec]	ATS Protocol with 8 Maneuvers & No Censor [L/sec]
Maximum Max	2.78	3.00	3.20
Average Max	2.53	2.60	2.64
Minimum Max	2.20	2.29	2.44
St Dev	0.166	0.166	0.196
CV_m	---	---	2.9%
CV_t	6.6%	6.4%	7.4%

Simulation Parameters

- Distribution (normal, from Brändli et al)
- Expected maximum FEV_1 (from Brändli et al)
- Inter-test coefficient of variation [CV_t] (from Künzli et al)
- Inter-maneuver coefficient of variation [CV_m] (from exploratory data collection)

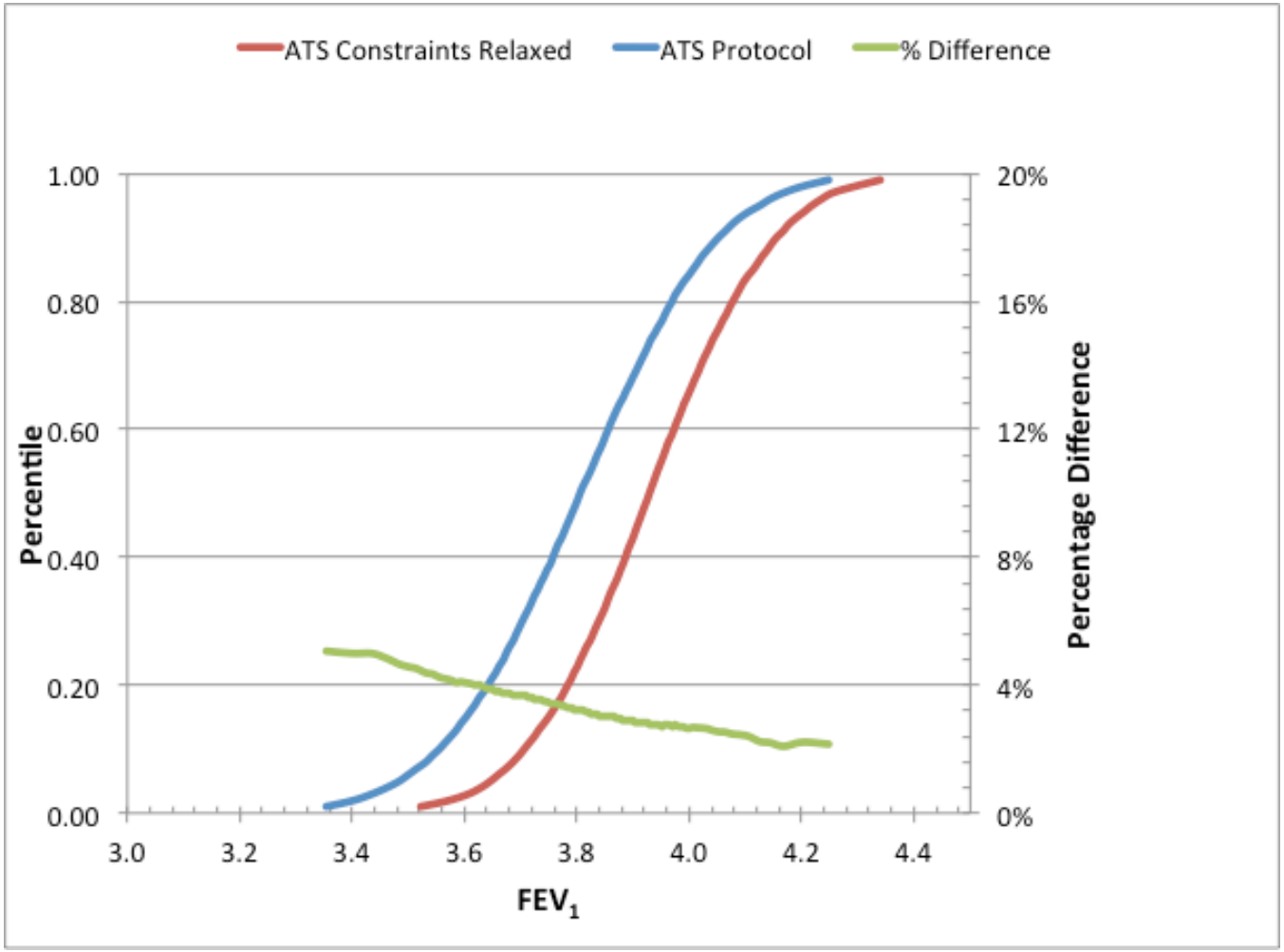
Baseline Simulation Model

- Non-smoking 5'8" 59-year old male subject
- 8 maneuvers/test, 10k tests
- User-adjustable default maximum FEV_1
- User-adjustable CV_t and CV_m

2 Stage Simulation		Model	Coefficient of Variation	StDev of Maxima
1	Across 10k tests	X_t = NORMDIST (RAND(), 3.72, σ_t)	CV_t = 4%	σ_t = .15
2	Across 8 maneuvers per test	X_m^t = NORMDIST (RAND(), X_t, σ_m)	CV_m = 4%	σ_m = .15

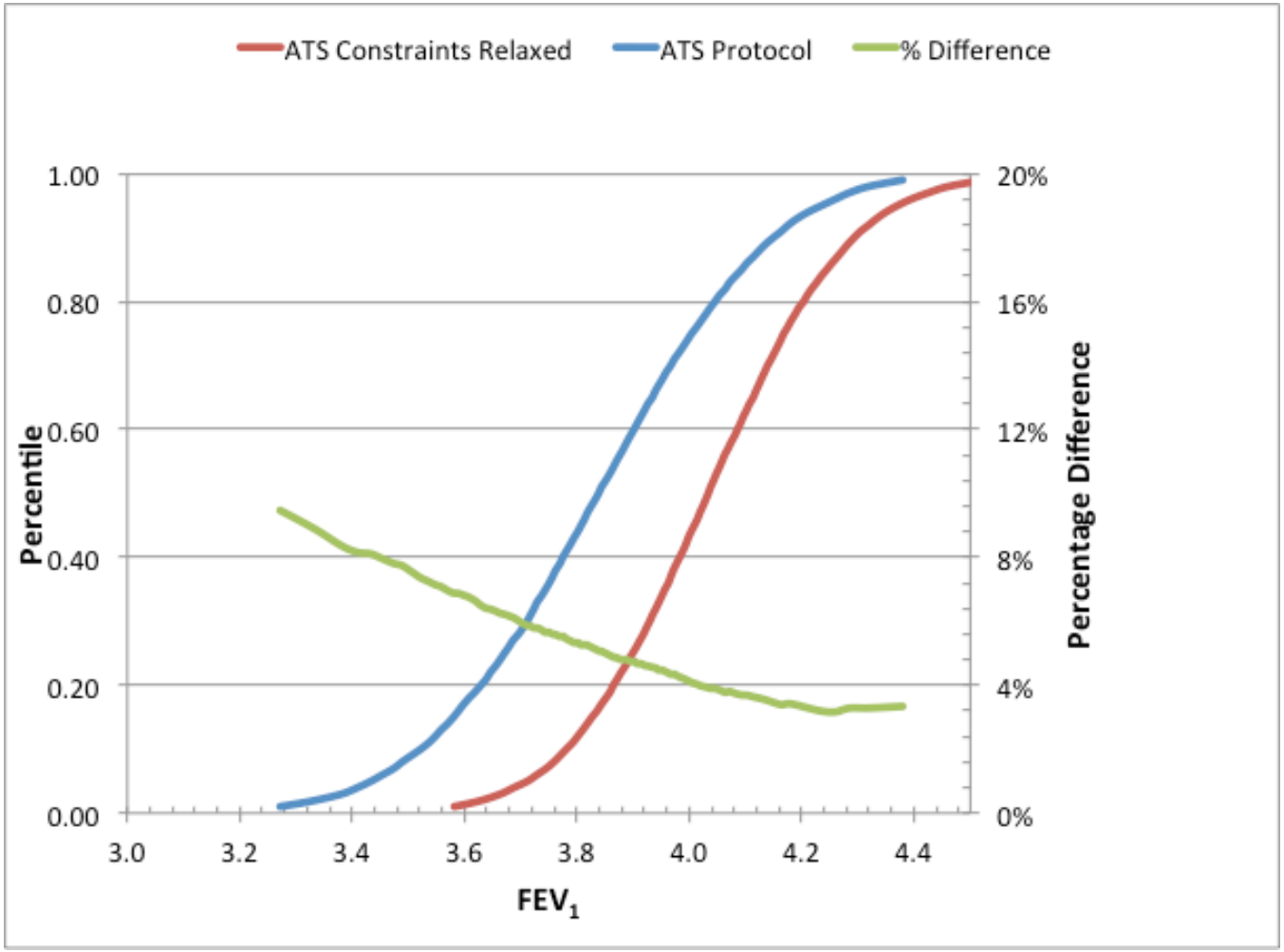
ATS Protocol v. Full Data Set

CV_m = 4%, CV_t = 4%



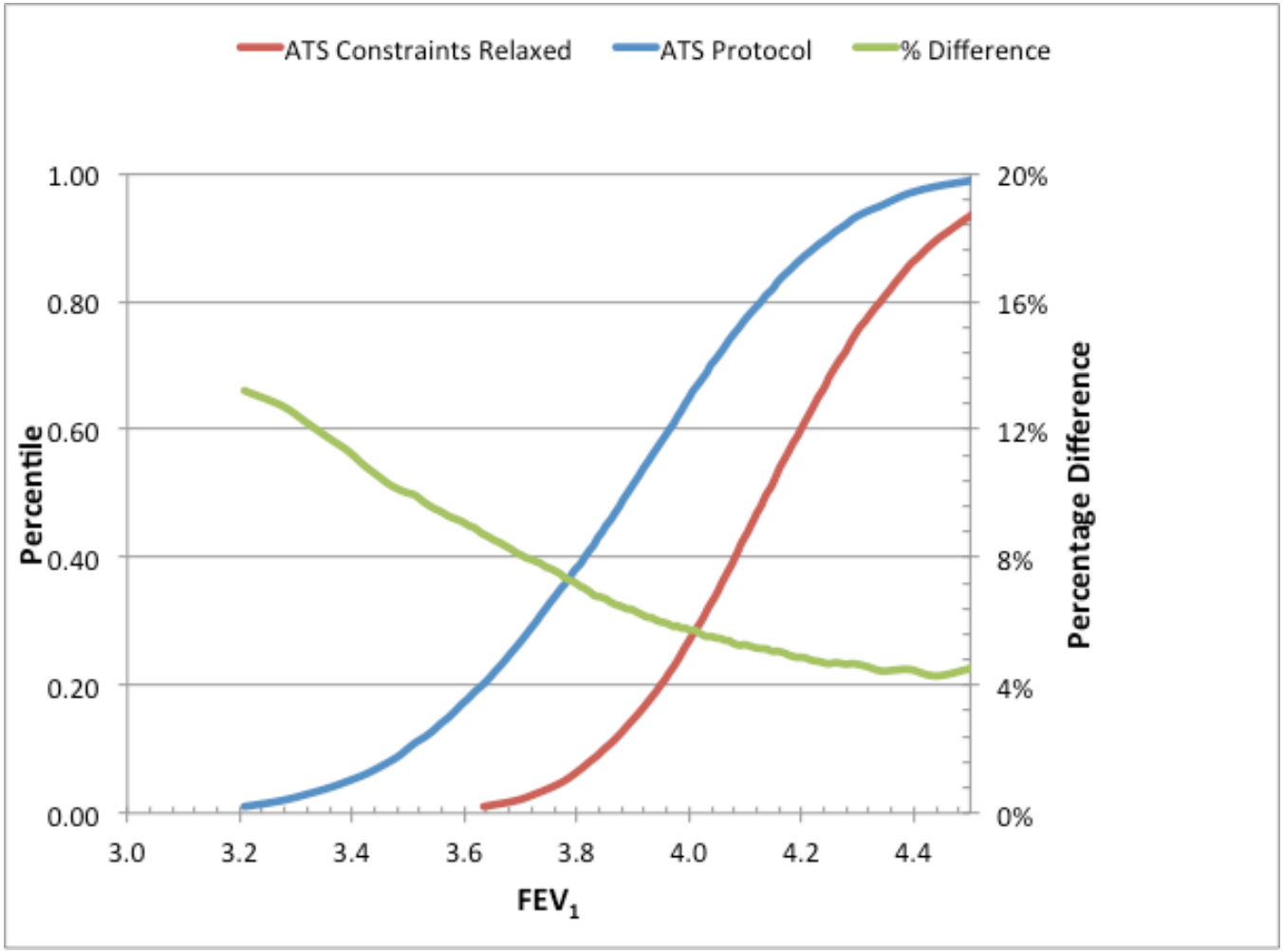
ATS Protocol v. Full Data Set

CV_m = 6%, CV_t = 4%



ATS Protocol v. Full Data Set

$CV_m = 8\%, CV_t = 4\%$



CONCLUSIONS AND NEXT STEPS

About ATS Protocol in Clinical Practice

- ATS stopping rule and censor result in unreported variance
  - Relaxing stopping rule and censor increase maximum FEV_1 and FVC
  - Amount of increase depends on stopping point, level of censor and CV_m and CV_t
- Therefore
  - ATS results may be arbitrary
  - Clinical classifications based on ATS results may be misleading

About ATS Protocol in Air Pollution Research

- Problems with ATS protocol are magnified because research focus is on small differences
- Small differences that are deemed biologically meaningful may overstate risk
  - artifacts of ATS protocol
  - due to chance
  - misinterpreted as statistically significant because CV_m and CV_t are assumed to equal zero

Next Steps (1)

- Publish
- Estimate likelihood of observing *x*% differences by chance as CV_m and CV_t vary
- Estimate minimum difference in concentration needed to infer that *x*% difference is not due to chance as CV_m and CV_t vary
- Review published literature to determine which studies have conclusions that may warrant reconsideration

Next Steps (2)

- Confirm our simulation results
- Revise air pollution research protocol to include 8 maneuvers with constant recovery time and no data censoring
- Estimate CV_m and CV_t for population and all test subpopulations
- Account for CV_m and CV_t in all future statistical analyses of air pollution effects