National Ambient Air Quality Standards Causal Framework

Julie E. Goodman, PhD, DABT, ACE, ATS

Ozone Webinar

Texas Commission on Environmental Quality

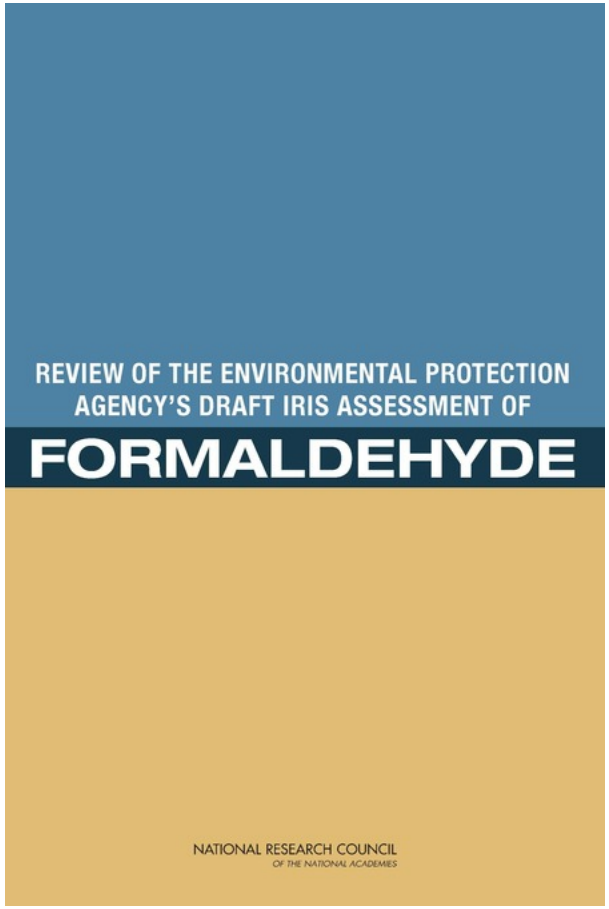
March 16, 2015

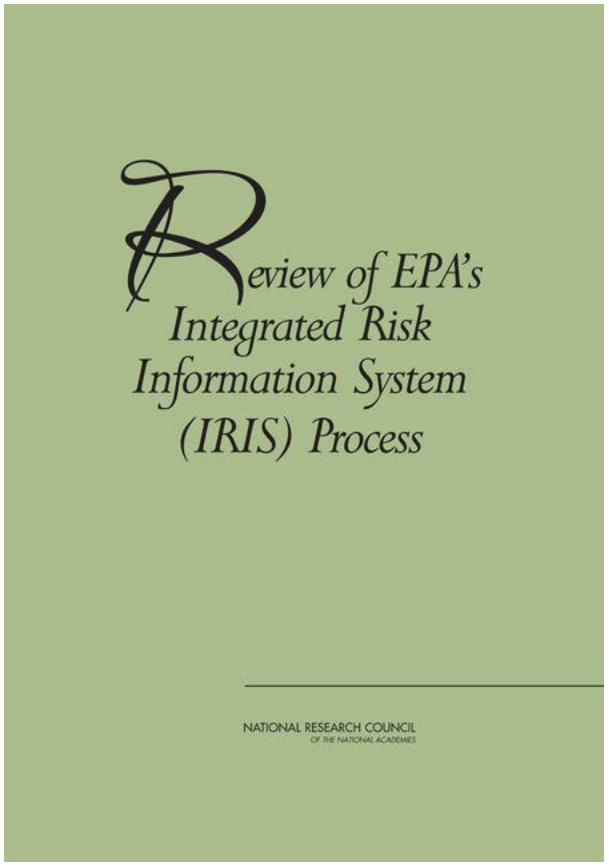
GRADIENT

NAAQS Process for Causal Determination

Step	Description
1	Literature search
2	Selection of studies for inclusion
3	Consideration of general limitations of each study type
4	Use of modified Bradford Hill aspects to aid in judging causality
5	Evaluate evidence for major health outcome categories
6	Integrate evidence from across disciplines and across health endpoints
7	Weight alternative views on controversial issues
8	Characterize strength of evidence into casual conclusions
9	Assess adversity of effects

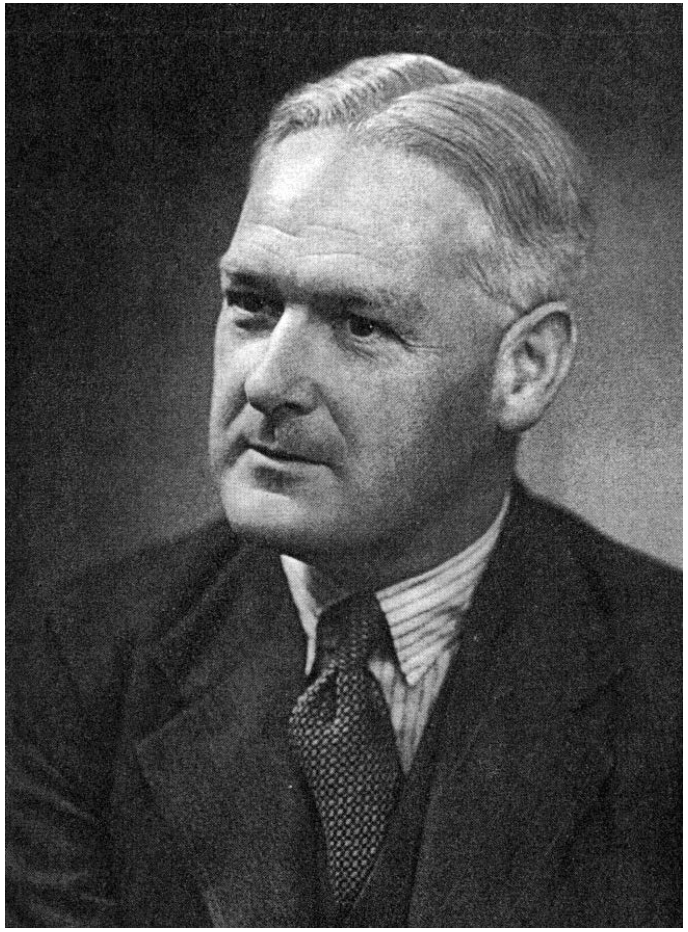
Recommendations for Evidence Integration





Bradford Hill "Criteria"

- Strength of association
- Consistency
- Specificity
- Temporality
- Exposure-Response
- Biological plausibility
- Coherence
- Experiment
- Analogy



Austin Bradford Hill
1897 - 1991

What is this list?

Alternative Hypotheses

"[I]s there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect?"

- Bradford Hill

Key Questions

Based on observations, what hypothesized causal processes are necessary? Sufficient?

How do they generalize? What *other* manifestations should they have?

If hypothesis were wrong, how *else* would one explain the array of outcomes?

GRADIENT

Relative Credence in Competing "Accounts"

"Account" – an articulated set of proposed explanations for the set of observations

Findings Indicate Target-Population Risk	Findings Do Not Indicate Target-Population Risk
• reasoning why • how contradictions resolved • why assumptions reasonable	• reasoning why **not** • how findings are **otherwise** explained • why assumptions reasonable

Institute of Medicine (IOM) Framework

Classification	Description
Sufficient	The evidence is sufficient to conclude that a causal relationship exists.
Equipoise and Above	The evidence is sufficient to conclude that a causal relationship is **at least as likely as not**, but not sufficient to conclude that a causal relationship exists.
Below Equipoise	The evidence is **not sufficient** to conclude that a causal relationship is **at least as likely as not**, or is not sufficient to make a scientifically formed judgment.
Against	The evidence suggests the lack of a causal relationship.

Institute of Medicine (IOM). (2008). Improving the Presumptive Disability Decision-Making Process for Veterans. Committee on Evaluation of the Presumptive Disability Decision-Making Process for Veterans, Board on Military and Veterans Health. National Academies Press [Online]. 781p.

GRADIENT

IOM *vs.* NAAQS Framework

IOM Causation Categories	NAAQS Causal Framework
Sufficient	Causal
Equipoise and above	Likely causal
Below equipoise	Suggestive
Against	Inadequate
	Not likely causal

EPA United States Environmental Protection Agency

Weight of Evidence for Causal Determination

… doses or exposures generally w/in 1-2 orders of magnitude of recent concentrations

. . . chance, confounding, and other biases could be ruled out with reasonable confidence

. . . multiple high-quality studies by multiple research groups

Table II Weight of evidence for causal determination.

	Health Effects	Ecological and Welfare Effects
Causal relationship	Evidence is sufficient to conclude that there is a causal relationship with relevant pollutant exposures (e.g., doses or exposures generally within one to two orders of magnitude of recent concentrations). That is, the pollutant has been shown to result in health effects in studies in which chance, confounding, and other biases could be ruled out with reasonable confidence. For example: (1) controlled human exposure studies that demonstrate consistent effects; or (2) observational studies that cannot be explained by plausible alternatives or that are supported by other lines of evidence (e.g., animal studies or mode of action information). Generally, the determination is based on multiple high-quality studies conducted by multiple research groups.	
Likely to be a causal relationship	Evidence is sufficient to conclude that a causal relationship is likely to exist with relevant pollutant exposures. That is, the pollutant result in health effects in studies not explained by chance, confo biases, but uncertainties remain overall. For example: (1) observ an association, but copollutant to address and/or other lines of human exposure, animal, or mo information) are limited or incon toxicological evidence from mul different laboratories demonstrate or no human data are available. Generally, the determination is based on multiple high-quality studies.	Evidence is sufficient to conclude that there is a likely causal association with relevant pollutant
Suggestive, but not sufficient, to infer a causal relationship	Evidence is suggestive of a causal relationship with relevant pollutant exposures but is limited, and chance, confounding and other biases cannot be ruled out. For example, (1) when the body of evidence is relatively small, at least one high-quality epidemiologic study shows an association with a given health outcome and/or at least one high-quality toxicological study shows effects relevant to humans in animal species; or (2) when the body of evidence is relatively large, evidence from studies of varying quality is generally supportive but not entirely consistent, and there may be coherence across lines of evidence (e.g., animal studies or mode of action information) to support the determination.	Evidence is suggestive of a causal relationship with relevant pollutant exposures, but chance, confounding, and other biases cannot be ruled out. For example, at least one high-quality study shows an effect, but the results of other studies are inconsistent.
Inadequate to infer a causal relationship	Evidence is inadequate to determine that a causal relationship exists with relevant pollutant exposures. The available studies are of insufficient quantity, quality, consistency, or statistical power to permit a conclusion regarding the presence or absence of an effect.	The available studies are of insufficient quality, consistency, or statistical power to permit a conclusion regarding the presence or absence of an effect.
Not likely to be a causal relationship	Evidence indicates there is no causal relationship with relevant pollutant exposures. Several adequate studies, covering the full range of levels of exposure that human beings are known to encounter and considering at-risk populations and lifestages, are mutually consistent in not showing an effect at any level of exposure.	Several adequate studies, examining relationships with relevant exposures, are consistent in failing to show an effect at any level of exposure.

3

Causal Relationship

- Evidence is sufficient to conclude that there is a causal relationship with relevant pollutant exposures (i.e., doses or exposures generally **within one to two orders of magnitude** of current levels).
- That is, the pollutant has been shown to result in health effects in studies in which **chance, bias, and confounding could be ruled out with reasonable confidence**. For example:
  - a) controlled human exposure studies that demonstrate consistent effects; or
  - b) observational studies that cannot be explained by plausible alternatives or are supported by other lines of evidence (e.g., animal studies or mode of action information).
- Evidence includes multiple high-quality studies

Likely to Be a Causal Relationship

- Evidence is sufficient to conclude that a causal relationship is likely to exist with relevant pollutant exposures, but **important uncertainties remain**.
- That is, the pollutant has been shown to result in health effects in studies in which **chance and bias can be ruled out with reasonable confidence but potential issues remain**. For example:
  - a) observational studies show an association, but copollutant exposures are difficult to address and/or other lines of evidence (controlled human exposure, animal, or mode of action information) are limited or inconsistent; or
  - b) animal toxicological evidence from multiple studies from different laboratories that demonstrate effects, but limited or no human data are available.
- Evidence generally includes multiple high-quality studies.

Suggestive of a Causal Relationship

- Evidence is suggestive of a causal relationship with relevant pollutant exposures, but is **limited**.
- For example,
  - (a) at least **one** high-quality epidemiologic study shows an association with a given health outcome but the results of other studies are inconsistent; or
  - (b) a well-conducted toxicological study, such as those conducted in the National Toxicology Program (NTP), shows effects in animal species.

Inadequate to Infer a Causal Relationship

- Evidence is inadequate to determine that a causal relationship exists with relevant pollutant exposures.
- The available studies are of **insufficient quantity, quality, consistency, or statistical power** to permit a conclusion regarding the presence or absence of an effect.

Not Likely to Be a Causal Relationship

- Evidence is suggestive of no causal relationship with relevant pollutant exposures.
- **Several adequate studies**, covering the full range of levels of exposure that human beings are known to encounter and considering at-risk populations, are **mutually consistent** in not showing an effect at any level of exposure.

IOM *vs.* NAAQS Framework

IOM Causation Categories
Sufficient
Equipoise and above
Below equipoise
Against

At-Risk Causation Categories
Adequate evidence
Suggestive evidence
Inadequate evidence
Evidence of no effect

NAAQS Causal Framework
Causal
Likely causal
Suggestive
Inadequate
Not likely causal

O_3 ISA Table 8-1 Classification of At-Risk Factors – Adequate Evidence

- There is **substantial, consistent evidence within a discipline** to conclude that a factor results in a population or lifestage being at increased risk of air pollutant-related health effect(s) relative to some reference population or lifestage.
- Where applicable this includes **coherence across disciplines**.
- Evidence includes multiple high-quality studies.

O_3 ISA Table 8-1 Classification of At-Risk Factors – Suggestive Evidence

- The collective evidence suggests that a factor results in a population or lifestage being at increased risk of an air pollutant-related health effect relative to some reference population or lifestage, but
- The evidence is limited due to **some inconsistency within a discipline** or, where applicable, **a lack of coherence across disciplines**.

O_3 ISA Table 8-1 Classification of At-Risk Factors – Inadequate Evidence

- The collective evidence is inadequate to determine if a factor results in a population or lifestage being at increased risk of an air pollutant-related health effect relative to some reference population or lifestage.
- The available studies are **of insufficient quantity, quality, consistency and/or statistical power** to permit a conclusion to be drawn.

O_3 ISA Table 8-1 Classification of At-Risk Factors – No Evidence of Effect

- There is **substantial, consistent evidence within a discipline** to conclude that a factor does not result in a population or lifestage being at increased risk of air pollutant-related health effect(s) relative to some reference population or lifestage.
- Where applicable this includes **coherence across disciplines**.
- Evidence includes multiple high-quality studies.

IOM *vs.* NAAQS Framework

IOM Causation Categories
Sufficient
Equipoise and above
Below equipoise
Against

At-Risk Causation Categories
Adequate evidence
Suggestive evidence
Inadequate evidence
Evidence of no effect

NAAQS Causal Framework
Causal
Likely causal
Suggestive
Inadequate
Not likely causal

NO_x ISA Table 5-1. Summary and description of scientific considerations for evaluating the quality of studies on the health effects from NO_x

Study Design		
Controlled Human Exposure	**Animal Toxicology**	**Epidemiology**
• Clearly defined hypotheses/aims • Appropriately matched control exposures • Randomization and allocation concealment • Balanced crossover (repeated measures) or parallel design studies	• Clearly defined hypotheses/aims • Appropriately matched control exposures • Randomization and allocation concealment • All groups handled and cared for equally	• Clearly defined hypotheses/aims • Key designs for short-term exposure: time series, case crossover, panel • Key designs for long-term exposure: prospective cohort, nested case-control • High power studies key: large sample sizes, multiple years, multicity studies

NO_x ISA Table 5-1 (cont'd)

Controlled Human Exposure:
Studies should clearly describe the primary and any secondary objectives of the study, or specific hypotheses being tested. Study subjects should be randomly exposed without knowledge of the exposure condition. Preference is given to balanced crossover (repeated measures) or parallel design studies which include control exposures (e.g., to clean filtered air). In crossover studies, a sufficient and specified time between exposure days should be provided to avoid carry over effects from prior exposure days. In parallel design studies, all arms should be matched for individual characteristics such as age, sex, race, anthropometric properties, and health status. Similarly, in studies evaluating effects of disease, appropriately matched healthy controls are desired for interpretative purposes.
Animal Toxicology:
Studies should clearly describe the primary and any secondary objectives of the study, or specific hypotheses being tested. Studies should include appropriately matched control exposures (e.g., to clean filtered air, time matched). Studies should use methods to limit differences in baseline characteristics of control and exposure groups. Studies should randomize assignment to exposure groups and where possible conceal allocation to research personnel. Groups should be subjected to identical experimental procedures and conditions and care of animals, including housing, husbandry, etc. Blinding of research personnel to study group may not be possible due to animal welfare and experimental considerations; however, differences in the monitoring or handling of animals in all groups by research personnel should be minimized.
Epidemiology:
Studies should clearly describe the primary and any secondary aims of the study, or specific hypotheses being tested. For short-term exposure, time-series, case crossover, and panel studies are emphasized over cross-sectional studies because they examine temporal correlations and are less prone to confounding by factors that differ between individuals (e.g., SES, age). Studies with large sample sizes and conducted over multiple years are considered to produce more reliable results. If other quality parameters are equal, multicity studies carry more weight than single-city studies because they tend to have larger sample sizes and lower potential for publication bias. For long-term exposure, inference is considered to be stronger for prospective cohort studies and case-control studies nested within a cohort (e.g., for rare diseases) than cross-sectional, other case-control, or ecologic studies. Cohort studies can better inform the temporality of exposure and effect. Other designs can have uncertainty also related to the appropriateness of the control group or validity of inference about individuals from group-level data. Study design limitations can bias health effect associations in either direction.

NO_x ISA Table 5-1 (cont'd)

Study Population/Test Model		
Controlled Human Exposure	**Animal Toxicology**	**Epidemiology**
• Similarly matched control and exposed subjects • Subject characteristics reported • Clearly indicated inclusion and exclusion criteria • Independent, clinical assessment of the health condition • Loss or withdrawal of subjects should be reported with rationale	• Animal characteristics reported • Studies testing and reporting both sexes and multiple life stages preferred • Loss or exclusion of animals should be reported with rationale	• Representative of population of interest • High participation and low drop-out over time that is not dependent on exposure or health status • Clearly indicated inclusion and exclusion criteria • Independent, clinical assessment of health condition • Groups are compared if from same source population

Pollutant		
Controlled Human Exposure	**Animal Toxicology**	**Epidemiology**
• Studies of NO_2 are emphasized	• Studies of NO_2 are emphasized	• NO_2 emphasized over NO, NO_X • Comparisons of health effect associations among gaseous oxides of nitrogen species ideal

NO_x ISA Table 5-1 (cont'd)

Exposure Assessment or Assignment		
Controlled Human Exposure	**Animal Toxicology**	**Epidemiology**
• Well characterized and reported exposure conditions • Limited to studies that utilize NO_2 and/or NO concentrations less than or equal to 5,000 ppb • Preference is given to studies that include exposure control groups • Randomized exposure groups	• Well characterized and reported exposure conditions • Inhalation exposure • Limited to studies that utilize NO_2 and/or NO concentrations less than or equal to 5,000 ppb • All studies should include exposure control groups • Randomized exposure groups	• Exposure metrics that accurately represent temporal or spatial variability for study area • Comparisons of exposure measurement methods • Indoor and total personal exposures can inform independent effects of NO_2 • Lag/duration of exposure metric correspond with time course for health effect

NO_x ISA Table 5-1 (cont'd)

Outcome Assessment/Evaluation		
Controlled Human Exposure	**Animal Toxicology**	**Epidemiology**
• Same manner of outcome assessment for all groups • Validated, reliable methods • Reporting of outcome assessment details • Blinding of endpoint evaluators • Appropriate timing of endpoint evaluation	• Same manner of outcome assessment for all groups • Validated, reliable methods • Reporting of outcome assessment details • Blinding of endpoint evaluators • Appropriate timing of endpoint evaluation	• Same manner of outcome assessment for all groups • Validated, reliable methods • Assessment is blind to exposure status • Appropriate timing of endpoint evaluation

Potential Copollutant Confounding		
Controlled Human Exposure	**Animal Toxicology**	**Epidemiology**
• Well-characterized exposure	• Well-characterized exposure	• Traffic-related copollutants are key: CO, $PM_{2.5}$, BC/EC, OC, UFP, metal PM components, VOCs • Also considered: PM_{10}, SO_2, O_3

NO_x ISA Table 5-1 (cont'd)

Other Potential Confounding Factors[e]		
Controlled Human Exposure	**Animal Toxicology**	**Epidemiology**
• Preference given to studies with adequate control of factors influencing health response	• Preference given to studies with adequate control of factors influencing health response	• Potential confounders related to health effect and correlated with oxides of nitrogen should be examined • Potential confounders vary by study design (temporally vs. spatially correlated) and by health effects

Statistical Methodology		
Controlled Human Exposure	**Animal Toxicology**	**Epidemiology**
• Clearly described and appropriate statistical methods for the study design and research question • Preference given to adequately powered studies • Consideration given to trends in data and reproducibility	• Clearly described and appropriate statistical methods for the study design and research question • Preference given to adequately powered studies • Consideration given to trends in data and reproducibility	• Multivariable regression adjusting for potential confounders ideal • Exception is multipollutant models. Multicollinearity can produce unreliable results • Results based on small sample sizes can be unreliable

Evidence Integration Frameworks

Critical Reviews in Toxicology

http://informahealthcare.com/txc
ISSN: 1040-8444 (print), 1547-6898 (electronic)

Crit Rev Toxicol, 2013; 43(9): 753–784
 DOI: 10.3109/10408444.2013.832727

informa healthcare

REVIEW

A survey of frameworks for best practices in weight-of-evidence analyses

Lorenz R. Rhomberg[1], Julie E. Goodman[1], Lisa A. Bailey[1], Robyn L. Prueitt[1], Nancy B. Beck[2], Christopher Bevan[3], Michael Honeycutt[4], Norbert E. Kaminski[5], Greg Paoli[6], Lynn H. Pottenger[7], Roberta W. Scherer[8], Kimberly C. Wise[2], and Richard A. Becker[2]

Critical Reviews in Toxicology

http://informahealthcare.com/txc
ISSN: 1040-8444 (print), 1547-6898 (electronic)

Crit Rev Toxicol, Early Online: 1–21
 DOI: 10.3109/10408444.2013.837864

informa healthcare

REVIEW ARTICLE

Evaluation of the causal framework used for setting National Ambient Air Quality Standards

Julie E. Goodman, Robyn L. Prueitt, Sonja N. Sax, Lisa A. Bailey, and Lorenz R. Rhomberg

Frameworks for Evaluating Study Quality

Study Quality Criteria System	Human Studies	Animal Studies	*In vitro* Studies
Integrated Risk Information System (IRIS) Risk of Bias (RoB) Evaluation	√	√	
The National Toxicology Program's (NTP's) Office of Health Assessment and Translation (OHAT) Approach	√	√	
Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) System	√		
Money *et al*. (2013) Approach	√		
Navigation Guide	√	√	
Animal Research: Reporting of *In vivo* Experiments (ARRIVE) Guidelines		√	√
Klimisch System		√	√
Organisation for Economic Co-operation and Development (OECD) Guidance Document (GD) 34		√	√
Toxicological Data Reliability Assessment Tool (ToxRTool)		√	√
Assessment of Multiple Systematic Reviews (AMSTAR) System			

Specified Study Quality Criteria

Criteria	Human (5 Frameworks)			Animal (7 Frameworks)			*In vitro* (4 Frameworks)		
	Report	Score	Y/N[a]	Report	Score	Y/N[a]	Report	Score	Y/N[a]
Study Objectives	1			3			2		
Study Design/Setting	3			5			2		
Participant/Animal Characteristics	3			6	1		–	–	--
Study Size	3			5	1	1	1		1
Study Power Analysis	1								
Blinding and Randomization		3		2	3		2		
Comparison/Control Groups	1	2		6	1		3	1	
Husbandry	--	--	–	3	3	1	–	–	--
Inclusion/Exclusion Criteria	1			2	2		–	–	--
Experimental Procedure	--	--	–	4	1		2	1	1
Participation Rate/Attrition	1	2		1	2		–	–	--
Statistical Methods	1	2	1	2	2	1	2		1
Exposure Measurement Methods/Dose Admin.	1	3	1	5	2		3	1	
Confounding and Bias	1	3	1		2		–	–	--
Outcome Assessment	1	2	1	3	2	1	2	1	1
Result Reporting	2	3		3	3	1	2		1
Adherence to Protocol, Deviations		2			2				
Limitations	2			2			2		
Interpretation and Implications	1		1	2			2		
Generalizability	2			3			3		
Funding Source/Conflict of Interest	3	1		3	1		1		

GRADIENT

Questions?

Julie E. Goodman, PhD, DABT, ACE, ATS
Gradient
jgoodman@gradientcorp.com