Development of a Novel Evidence Integration Framework that Incorporates Human, Animal and Mechanistic Data

AUTHORS

Sandra Sulsky, Tracy Greene, Greg Mariano, Farah Chowdhury, Allison Franzen, P. Robinan Gentry, Ramboll; Donna C. Smith, Willie J. McKinney, Jr., Altria Client Services

ABSTRACT

To accurately characterize human health hazards, human, animal and mechanistic data must be integrated and the relevance to the research question of all 3 lines of evidence must be considered. Mechanistic data are often critical in fully integrating animal and human data and characterizing relevance and uncertainty. This novel evidence integration framework (EIF) provides a method for synthesizing data from comprehensive, systematic, quality-based assessments of the epidemiological and toxicological literature, including *in vivo* and *in vitro* mechanistic studies. The data are organized using both a disease-based and mechanism-based scheme, providing a method for assimilating and using mechanistic information to support evidence synthesis. The disease-based scheme uses the evidence of human health outcomes studied in the best quality epidemiological literature to organize the toxicological data according to authors' stated purpose, with the pathophysiology of the disease determining the potential relevance of the toxicological data. The mechanism-based scheme organizes the data based on the proposed mechanisms of effect and mechanistic data supporting key events leading to each human health endpoint, with the epidemiological data providing corroboration or no corroboration of causality. The EIF includes a method to cross-classify and describe the concordance of the data, and to characterize its uncertainty. A case study with nicotine is presented focused on the integration of evidence related to non-acute exposure and cancer. The results of the case study highlight knowledge gaps, demonstrate how different conclusions may be drawn depending on the organiz ation of the data, and show the impact of uncertainties on the strength of causal inference.

INTRODUCTION

Scientifically justifiable decisions regarding the potential for health effects resulting from chemical exposure require integration of multiple streams of evidence (i.e., epidemiological, animal and mechanistic), as well as consideration of the breadth and quality of that evidence.

METHODS

The EIF is applied following a comprehensive, systematic, critical review of the epidemiological and toxicological literature. This case study demonstrates the EIF by applying it to assess the relationship between non-acute nicotine exposure and cancer. The epidemiological literature focused on comparisons between users of smokeless tobacco – as a surrogate for nicotine exposure – compared with non-users of tobacco. The toxicological data were for chronic or repeated *in vivo* or *in vitro* exposure to nicotine, *per se*.

The best quality epidemiological and toxicological literature was identified and documented using standard techniques (Table 1 and Table 2), with modifications as necessary to allow for the evaluation of quality for mechanistic data.

Table 1. Epidemiology study attributes and their contributions to assessment of quality[1]

Study attribute	Quality criteria
Study objectives	Clearly stated Relevant to research questions
Study methods	Adequately described Appropriate for objectives Minimize selection and information bias Reasonable statistical power
Outcome measurement	Well-defined, reasonably specific Accurate measurement or diagnosis Proper time frame for risk of outcome
Exposure measurement	Well-defined, specific Verified (e.g., with biomarkers) Accounts for changes over time
Control of confounding	Known risk factors considered and measured Reasonable analysis method(s) used (stratification, multivariate statistical models)

Table 2. Toxicology study attributes and their contributions to assessment of quality[2]

Study quality category	Quality score	Requirements
Adequate	1	Studies or data from the literature or reports that are carried out or generated according to generally valid and/or internationally accepted testing guidelines (preferably performed according to good laboratory practices [GLP]), or studies in which the test parameters documented are based on a specific (national) testing guideline (preferably performed under GLP), or studies in which all parameters described are closely related/comparable to a guideline method.
	2	Studies that are not performed according to GLP or specific testing guidelines but are well documented and scientifically acceptable For *in vivo* animal studies this includes the following criteria: • Data on the test animals including species, sex and strain • Purity/composition/origin of the test substance • Number of animals evaluated • Scope of the investigation per animal (e.g., clinical chemistry, hematology, organ weights, pathology or histopathology) and description of the methods • Description of the changes observed • Control group or historical control data of the laboratory • Description of the test conditions • Description of the route and doses of administration • Dose/concentration relationship is possible For *in vitro* studies this includes the following criteria: • Description of the test system and test method in details • Purity/composition/origin of the test substance (for *in vitro* studies only, if the suppler was noted it was assumed that the purity was acceptable) • Data on the dose/concentration differentiated according to the toxicity of the test substance on the test system; information on volatility • Data on secondary effects which may influence a result (solubility, impurities, pH shifts, influence on the osmolarity, etc) • Appropriate negative/positive controls as integral parts of the test References on adequacy of the method should be given or generally known.
Fair	3	Studies or data from the literature which do not meet the criteria for a quality score of 2. These include studies for which the methods were not clearly defined or specified in the study, in which only one dose was administered, or small numbers of animals per group were tested. The appropriate number of animals per group was determined based on recommendations from OECD guidelines for comparable study types.
	4	Studies or data from the literature which do not provide sufficient experiment details and which are only listed in short abstracts or secondary literature (books, reviews, etc).
Not useful	5	Criteria for *in vivo* animal studies: • Exposure not specific to chemical of interest • Single or acute exposure • Inappropriate or irrelevant species tested • Route of exposure not relevant to humans • If the endpoint being considered was related to addiction or reproductive/developmental toxicity • If the animals tested underwent any type of alteration or injury prior to testing Criteria for *in vitro* studies: • Exposure not specific to the chemical of interest • The endpoint was not related to the human health outcome(s) of interest, according to the author's purpose

After synthesizing each line of evidence independently, the data are integrated in stages and a gap analysis can be performed.

Stage 1: Disease-based integration

After identifying the human health endpoints that have been evaluated in the best-quality epidemiological studies, the toxicological evidence pertinent to each endpoint is organized in 3 ways, based on the proposed pathophysiology (i.e., the data must be demonstrated to be relevant to a key event in the development of the disease).

- The data are grouped to correspond to the human health outcomes from the epidemiological literature and divided into those providing evidence of a statistically significant change in response to exposure (evidence of an effect) and those that do not (no evidence of an effect)
- The relevance of the toxicological data to a human health outcome is classified as likely or uncertain based on what is known regarding the modes of action of the chemical relative to the development and/or progression of the specific human disease
- The toxicological data are classified into 1 of 6 major and minor domains (i.e., pathology, gross, functional, biochemical, genotoxicity and gene expression)[3]
  - Major evidence: endpoints clearly related to or associated with the development of the human disease, such as the apical endpoint; gross physiological or histopathological changes that are clearly analogous to health effects observed in humans with relevant exposures; or that have been identified as key events in the development of specific human health effects
  - Minor evidence includes endpoints that are less directly analogous or are less clearly related to the human disease in question
- Conclusions based on the integrated toxicological data are shown in Table 3

Table 3. Disease-based integration of overall toxicological conclusions:[a]

		If the majority of evidence in minor domains is:[a]								
		Insufficient	**Balanced**[b]		**Suggestive**		**Not suggestive**	
If the majority of the evidence in major domains is:[a]	**Insufficient**	a. Insufficient	b. Balanced[c]	c. Insufficient[d]	d. Suggestive[c]	e. Insufficient[e]	f. Not suggestive[c]	g. Insufficient[e]
	Balanced[b]	h. Balanced	i. Balanced		j. Suggestive[c]	k. Balanced[d]	l. Not suggestive[c]	m. Balanced[d]
	Suggestive	n. Suggestive	o. Suggestive		p. Suggestive		q. Balanced[c]	r. Suggestive[d]
	Not suggestive	s. Not suggestive	t. Not suggestive		u. Suggestive[c]	v. Not suggestive[d]	w. Not suggestive	

[a] The major and minor domains are: biochemical, pathological, functional, gross or genotoxic. The gene expression domain is minor, only (no corresponding major domain)

[b] To account for uncertainties, "balanced" means 40–60% of the endpoints suggest evidence of an effect from nicotine exposure

[c] Greater than 60% of the minor evidence is from a primary cell line, or from *in vivo* gross data related to precancerous lesions

[d] Greater than 60% of the minor evidence is from a cancer cell line, or from *in vivo* gross data related to the apical endpoint but not of relevant route or duration to observe endpoint, or from an animal model with questionable relevance to human health (e.g., AJ mouse model)

Stage 2: Mechanism-based integration

The mechanism-based component of the EIF is driven by the toxicological data. For this cancer analysis, mechanisms related to direct measurements of cancer hallmarks or enabling characteristics.[4] Endpoints are categorized as being of likely versus uncertain relevance based on their specificity. Epidemiological data corroborate toxicological findings based on similarity to apical endpoints or tissue-type concordance.

Stage 3: Overall synthesis

The overall conclusions from the epidemiological and toxicological data are integrated as shown in Table 4.

Table 4. Integrated summary conclusions possible based on the integration of the epidemiological and toxicological conclusions

		If the toxicology evidence is:								
		Insufficient	**Balanced**		**Suggestive**		**Not suggestive**	
If the epidemiology evidence is:	**Insufficient**[a]	a. Insufficient	b. Balanced[b]	c. Insufficient[b]	d. Suggestive		e. Insufficient[c]	f. Not suggestive[d]
	Balanced	g. Balanced	h. Balanced		i. Suggestive		j. Not suggestive[e]	k. Balanced[e]
	Suggestive[f]	l. Suggestive	m. Suggestive		n. Suggestive		o. Balanced[g]	p. Suggestive[g]
	Not suggestive	q. Not suggestive	r. Not suggestive		s. Suggestive[g]	t. Not suggestive[g]	u. Not suggestive	

[a] If there are no relevant endpoints available in the epidemiology data of adequate quality, include that statement and indicate the conclusions are based on toxicology data only

[b] If the toxicological evidence is balanced and is from minor domains and/or the majority of the data (i.e., > 60%) is from cancer a cell line, the overall evidence is insufficient. If the toxicological evidence is balanced and is from major domains and/ or the majority of the data (i.e., > 60%) is from primary cell lines, the overall evidence is balanced

[c] Insufficient amount of epidemiological data that is of adequate quality

[d] Insufficient amount of epidemiological data that is of adequate quality and toxicological data related to the human apical outcome is not suggestive of an effect

[e] If the toxicological evidence is not suggestive and is only from minor domains and/or the majority of the data (i.e., > 60%) is from cancer cell lines, the overall evidence is balanced or suggestive (depending on the overall epidemiological conclusion); if the toxicological evidence is not suggestive and is from major domains and primary cell line data, the overall conclusion is that the evidence is not suggestive

[f] If epidemiology data corroborate the toxicology data, then the conclusion is pushed up one level (e.g., from balanced to suggestive)

[g] If the toxicological evidence is suggestive and is from minor domains and/or the majority of the data (i.e., > 60%) is from a cancer cell line, the overall evidence is not suggestive. If the toxicological evidence is suggestive and is from major domains and/or the majority of the data (i.e., 60%) is from primary cell lines, the overall evidence is suggestive

RESULTS

Searches for studies of any type of cancer in smokeless tobacco users versus non-users and for cancer-related endpoints studied in connection with repeated exposure to nicotine returned 32,247 potentially relevant articles.

- 44 epidemiology studies of 48 types of cancer
- 314 toxicology studies; 100/314 relevant to the disease-based component; 312/314 relevant to the mechanism-based component (some studies provide data that are relevant to either or both components)

Disease-based component of the EIF: Two examples illustrating dependence of results on outcome definitions

The evidence for 22 specific cancers was evaluated based on signals indicating potential associations between smokeless tobacco use and various cancers available in the epidemiological evidence. There were limited epidemiology data for 26 other specific cancers, yielding insufficient evidence to apply the EIF. Examples of how the evidence was organized is provided in Figures 1 and 2, which assisted in drawing conclusions and integrating evidence using Table 6.

Figure 1. Stomach cancer

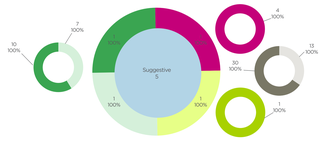


Figure 2. Gastrointestinal/digestive system cancer

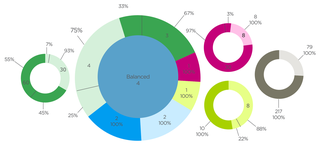
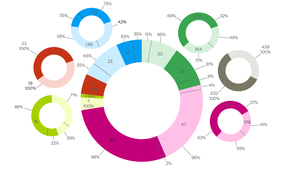


Figure 3. Mechanism-based approach



The conclusions drawn from the available data depend on the organizational scheme and outcome definitions. Table 5 summarizes the data for digestive systems cancers.

Table 5. Summary of epidemiological, toxicological, and overall integrated results for the disease-based approach to evidence integration

Cancer type	Epidemiological[a]	Toxicological[b]	Overall[c]
Digestive system			
Stomach cancer	Suggestive	Balanced	Suggestive
Colon cancer	Suggestive	Not suggestive	Not suggestive
Pancreatic cancer	Balanced	Balanced	Balanced
Rectal cancer	Balanced	Not suggestive	Not suggestive
Gastrointestinal/digestive system cancers combined	Balanced	Balanced	Balanced
Overall conclusion: The integrated toxicological and epidemiological evidence suggests that **nicotine may not play a role** in the development of **digestive system cancers.**			

[a] Integrated epidemiological evidence relating smokeless tobacco use to specified diseases shown in the table rows

[b] Integrated toxicological evidence relating nicotine exposure to the development of endpoints (not necessarily the apical endpoint) potentially associated with the development of the diseases shown in the table row

[c] The overall conclusions reflect the integration of the epidemiological and toxicological evidence according to the system summarized in Table 6

Mechanism-based component of the EIF

312 studies reported data potentially relevant to the hallmarks and enabling characteristics of cancer.[4] They reported 2,420 relevant endpoints which were categorized into the domains as providing either major or minor evidence (Figure 5).

Table 6. Summary and integration of evidence from the mechanism-based component of the EIF

Domain	Epidemiological	Toxicological	Overall
Gross	Balanced	Not suggestive	Not suggestive
Functional	Insufficient	Balanced	Balanced
Pathological	Suggestive	Balanced	Suggestive
Biochemical	Balanced	Balanced	Balanced
Genotoxicity	Balanced	Suggestive	Suggestive
Expression change	Insufficient	Balanced	Balanced

Gap analysis

- In the human domain there are no relevant studies of nicotine *per se*. Studies of smokeless tobacco users tend to be subject to exposure misclassification and to provide incomplete information on key details of exposure, including duration and amount of smokeless tobacco used prior to disease detection
- The toxicological data provide information about exposure to nicotine *per se*, but very few studies directly measured the incidence of cancer in animals *in vivo*. Such studies would provide the strongest evidence of a potential relationship between nicotine exposure and cancer risks in humans
- The limited number of repeated dose studies that focused on tumor incidence in animals *in vivo* could be viewed as a data gap, or as a consequence of the general movement away from animal testing. A more pertinent and general question to be addressed is whether *in vitro* gene expression studies that are being substituted for animal studies can predict the development of the disease of interest

Observations based on the examples

- Specificity in both the outcome and exposure measures is required to link the epidemiological with the toxicological data, and to draw appropriate discipline-specific conclusions
- The examples for nicotine illustrate how conclusions can be affected by the ways the endpoints are defined and combined
- The EIF provided a sensitive system for detecting signals potentially indicating a nicotine effect; these signals should be interpreted with full appreciation of the amount and quality of the available evidence and any uncertainties related to drawing conclusions (e.g., using smokeless tobacco as a surrogate for nicotine per se exposure) that can be used as guideposts for future, focused research and surveillance

CONCLUSIONS

- There are inherent differences in the data available from the human observational studies (e.g., the epidemiological evidence relates to tobacco, a mixed exposure) and the data available from experiments (e.g., the toxicological evidence that relates to nicotine exposure)
- The EIF accounts for quality and quantity of the underlying data, and uncertainties in the conclusions the EIF can support
- The EIF uses information about specific disease etiology to link the epidemiological and toxicological evidence
- The disease-based component identifies toxicological data relevant to human health endpoints via pathophysiological processes that lead to the disease
- The mechanism-based component directly assesses the evidence in major and minor domains that address the potential for exposure to affect the progression to a disease state or health endpoint
- The EIF is flexible and can incorporate new information at any level (epidemiological, *in vivo* and *in vitro*), as it becomes available
- The EIF systematizes integration of epidemiological and toxicological data derived from comprehensive, systematic, and critical reviews of the relevant literature. Applying the EIF can identify previously undetected signals of potential causal associations between chemical exposures and human health outcomes and identifies knowledge gaps. These results provide a direction for future, targeted investigations to address the uncertainties in the available existing evidence

REFERENCES

1. Sulsky, S. I., F. H. Hooven, et al. (2002). "Critical review of the epidemiological literature on the potential cardiovascular effects of occupational carbon disulfide exposure." Int Arch Occup Environ Health 75(6): 365-380.

2. Criteria modified from Klimisch, H. J., M. Andreae, et al. (1997). "A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data." Regul Toxicol Pharmacol 25(1): 1-5.3. R Core Team. R: A language and environment for statistical computing, 2015.

3. Building on Dekant, W. and J. Bridges (2016) "A quantitative weight of evidence methodology for the assessment of reproductive and developmental toxicity and its application for classification and labeling of chemicals." Regul Toxicol Pharmacol 82: 173-185

4. Hanahan D, Weinberg RA. (2011) Hallmarks of cancer: the next generation. Cell 144(5):646-74