

Chapter 8

Assessing Exposures



Abstract An important problem in causal inference in medicine involves establishing causal relationships between environmental exposures and negative health outcomes. It is typically not possible to use RCTs to solve this problem, for ethical reasons. The approach outlined in this book is compared to two other prominent approaches: the procedures of the International Agency for Research on Cancer (IARC), and SYRINA, a framework for detecting exposures that affect the endocrine system.

An important problem in causal inference in medicine involves establishing causal relationships between environmental exposures and negative health outcomes (Hill 1965). Experimental studies, e.g., randomized controlled trials, tend to provide relatively strong evidence for causal claims. However, when assessing exposures it is typically not possible to carry out such trials in human populations, because this would involve unethically intervening to expose individuals to factors that are suspected to have deleterious health effects. The only available epidemiological studies are observational. As a result, it is difficult to obtain epidemiological data that are sufficient to establish causality.

This problem occurs, for instance, when assessing whether an environmental exposure is carcinogenic in humans. In such cases, different types of evidence are required. For example, the International Agency for Research on Cancer (IARC) attempts to determine whether particular exposures cause cancer in humans by looking at a variety of different types of evidence, namely, epidemiological studies, studies in experimental animals, and mechanistic and other relevant data (IARC 2015). The problem also occurs in assessing whether an exposure is an endocrine disruptor. In this context, Vandenberg et al. (2016) introduced SYRINA, a framework for the systematic review and integrated assessment of exposures. In this chapter, we compare the approach to assessing exposures given in this book with these other prominent approaches. First compare our approach to external validity to the approach endorsed by IARC, with reference to the example of establishing carcinogenicity of

benzo[a]pyrene—a compound that IARC recently evaluated and decided to upgrade from probable human carcinogen to human carcinogen largely based on just the mechanistic evidence and evidence from cancer bioassays. We then compare our approach to SYRINA, a framework for detecting exposures that affect the endocrine system (Sect. 8.2).

8.1 Comparison to IARC

Here we compare our approach to external validity to that of the International Agency for Research on Cancer (IARC). A note on terminology here. IARC use the term *generalizability*, as well as external validity, and for the purpose of this discussion we will regard them as synonymous. First, consider an example:

Example: *Carcinogenicity of benzo[a]pyrene.*

Benzo[a]pyrene is a polycyclic aromatic hydrocarbon (PAH) that is formed during incomplete combustion of organic material. Benzo[a]pyrene and other PAHs are an important industrial pollutant in soil, water, air, and sediments. They are also found in high concentrations in tobacco smoke, and in some pharmaceutical products. Human exposure occurs mainly through industrial and environmental exposure (IARC 2009). IARC has evaluated benzo[a]pyrene in four monographs, and it is currently classified as Group 1, carcinogenic to humans (IARC 2015).

In the most recent evaluation, epidemiological data were not available to the IARC working group. The working group therefore made its decision to classify benzo[a]pyrene as carcinogenic to humans based on mechanistic evidence and evidence from experimental animals. This makes the case of benzo[a]pyrene especially interesting for our purposes, as according to the procedure outlined above in Sect. 7.2, the correlation between benzo[a]pyrene and cancer required to establish the causal claim in humans would have to be inferred from observed outcomes in the experimental animals together with the mechanistic data.

On the approach of this book, first one formulates the causal claim under scrutiny: here, ‘benzo[a]pyrene causes cancer in humans’. In the context of IARC, this is to be taken as a qualitative claim—IARC identifies cancer hazards, and the exact size of the effect by which exposure increases cancer risk does not play a role in determining carcinogenicity. We should note though that a qualitative understanding of effect size does play a role in determining carcinogenicity. The IARC process is explicitly based on the causal indicators set out by Hill (1965), as we discuss above.

Next, one should assess—according to a suitable framework—the evidence for a correlation between the exposure and its effect, and articulate any hypo-

thetical mechanisms that would account for the correlation. Note that IARC use their own framework for assessing correlations (IARC 2015). A GRADE-like framework would be potentially useful in this context too—assuming that suitable modifications can be made to allow for differences in the understanding of bias in evidence that are appropriate for this change in purpose.

The evidence for the relevant mechanisms should then be graded according to the procedures described in Chap. 6. In the latest IARC monograph on benzo[a]pyrene, all the evidence of a correlation between the exposure and cancer came from studies on experimental animals—no epidemiological data were evaluated. The correlation between exposure and cancer in humans must thus be inferred via extrapolation from corresponding data in the experimental animals. This is based on assessing the evidence for correlation in the experimental animals, and assessment of similarities of the underlying mechanisms. The IARC monograph reports evidence of cancer outcomes upon exposure to benzo[a]pyrene in experimental animals. This was judged to be of high quality, both in terms of the validity of the research within species of experimental animals, and in terms of the additional corroboration gained by these results being robust across eight species of experimental animals (IARC 2009, 112–131). In addition, evidence is presented and evaluated for two main types of mechanism by which benzo[a]pyrene causes DNA adducts to form at known cancer hotspots: in one of these a metabolite of benzo[a]pyrene binding the DNA molecule, and the other an oxidized form of benzo[a]pyrene. In addition, similar activity of benzo[a]pyrene is reported to be shown in *in vitro* studies on human cell lines (IARC 2009, 131–137).

IARC considered there to be sufficient evidence for carcinogenicity in the experimental animals, *i.e.*, the causal claim about the experimental animals was established. IARC's current practice is to make some evaluations about possible mechanisms of carcinogenesis using a set of key characteristics shown by carcinogens (Smith et al. 2016). This is broadly compatible with the approach of this book, as there is high quality evidence of both correlation and underlying mechanisms in the experimental animals. This alone would not suffice to transfer the same claim to humans (nor does the IARC approach consider this). However, strong evidence of similar mechanisms operating in the experimental animals and humans, and the robustness of the experimental animal results across many species, warrants a mechanism-based extrapolation of the causal claim from the experimental animals to humans (Wilde and Parkkinen 2017). This, together with the mechanistic evidence directly on humans, such as evidence of formation of DNA adducts, is what, on the approach presented here, warrants establishing a causal conclusion about humans. In mechanism-based extrapolation, one compares the mechanisms responsible for an outcome in the target—of which a conclusion about causality is to be made—and in the study—about which direct evidence of causality is available—and looks for differences that might lead to differences in the outcome of interest between the study and the target. Here the outcome of interest is the development of tumours or the appearance of various cancer biomarkers upon exposure to

benzo[a]pyrene. A dependence between these outcomes and benzo[a]pyrene has been robustly demonstrated in the experimental animals. The relevant mechanisms are the pathways by which benzo[a]pyrene causes DNA adducts that can trigger tumorigenesis, that would explain the dependence. For these, there is evidence from cultured human cell lines, as well as the experimental animals, demonstrating strong similarities, and no differences that would indicate that benzo[a]pyrene does not cause cancer in humans. In addition, there is concordant evidence of the outcomes in several species of experimental animal, lending further credibility to the assumption that the carcinogenicity of benzo[a]pyrene is not dependent on idiosyncratic features of any particular species. These considerations, taken together, suffice to establish the carcinogenicity of benzo[a]pyrene in humans.

While the approach of this book would yield the same conclusion as IARC's, it should be noted that the procedures differ at certain points. IARC does not formally endorse extrapolation from experimental animals. Note though that this does not preclude altogether judgements about possible carcinogens where no human research is available, as in cases where only animal studies are available substances may be classified by IARC as belonging to *Group 2B: The agent is possibly carcinogenic to humans*. Nor does IARC formally endorse robustness of evidence as grounds for upgrading a classification, but allows for upgrading (or downgrading) a classification of carcinogenicity on the basis of mechanistic evidence alone. On the approach of this book, one may appeal to the aforementioned considerations, and one needs in addition to establish correlation in humans (by direct observation or extrapolation), before any claim about causality can be considered established.

Having considered an example, we now compare the general approach of this book to external validity to that of IARC. IARC's approach is summarized in Fig. 8.1.

The categories of IARC roughly correspond to those presented here, as follows. IARC have a ranking for overall carcinogenicity:

- Group 1* : Established
- Group 2a* : Provisionally established
- Group 2b* : Arguably true
- Group 3* : Speculative
- Group 4* : Ruled out

IARC also has a separate ranking of evidence of carcinogenicity in humans and animals:

- Sufficient* : Established
- Limited* : Provisionally established
- Inadequate* : Arguable or speculative
- Evidence Suggesting Lack of Carcinogenicity (ESLC)* : Ruled out

		EVIDENCE IN EXPERIMENTAL ANIMALS			
		<i>Sufficient</i>	<i>Limited</i>	<i>Inadequate</i>	<i>ESLC</i>
<i>Sufficient</i>		Group 1			
<i>Limited</i>		↑1 strong evidence in exposed humans Group 2A	↑2A belongs to a mechanistic class where other members are classified in Groups 1 or 2A Group 2B (exceptionally, Group 2A)		
EVIDENCE IN HUMANS		↑1 strong evidence in exposed humans ↑2A strong evidence ... mechanism also operates in humans	↑2A belongs to a mechanistic class ↑2B with supporting evidence from mechanistic and other relevant data	↑2A belongs to a mechanistic class ↑2B with strong evidence from mechanistic and other relevant data	Group 3 ↓4 consistently and strongly supported by a broad range of mechanistic and other relevant data
<i>Inadequate</i>		↓3 strong evidence ... mechanism does not operate in humans	Group 3	Group 3	
<i>ESLC</i>		Group 3			Group 4

Fig. 8.1 IARC’s approach to classifying potential carcinogens (<http://monographs.iarc.fr/ENG/Publications/Evaluations.pdf>)

In addition, IARC has a separate ranking of evidence of mechanisms:

- Strong* : Established
- Moderate* : Provisionally established
- Weak* : Arguable or speculative

What is being assessed by these three categories is a general mechanistic claim: e.g., the existence of a mechanism of action in animals; or the similarity of mechanism of action in humans to that in animals; or the existence of a mechanism of action in humans.

The approach of this book is simpler than that of IARC in one respect: a single scale from established to ruled out, rather than three different categorisations. On the other hand, the scale adopted in this book involves more categories.

In order to compare the approach of this book with that of IARC, consider two tables that illustrate the approach that this book takes with respect to external validity. First, Table 8.1 assumes that causality in the study has been established and charts similarity of mechanisms in the study and target populations against causation in the target population on the basis of evidence obtained on the target population. A second table, Table 8.2, assumes that similarity of mechanism is established and charts causation in the study population against causation in the target population on the basis of evidence obtained in the target population.

Table 8.1 Determining the status of the causal claim from similarity of mechanisms in the study and target populations and causation in the target population on the basis of evidence obtained on the target population. It is assumed here that causality in the study population has been established

		Similarity of mechanism in target and study				
		Established	Provisionally established	Arguable	Speculative	Ruled out
Causation from target studies	Established	Established	Established	Established	Established	Established
	Provisionally established	Established	Provisionally established	Provisionally established	Provisionally established	Provisionally established
	Arguable	Established	Provisionally established	Arguable	Arguable	Arguable
	Speculative	Established	Arguable	Speculative	Speculative	Speculative
	Ruled out	Provisionally ruled out	Ruled out	Ruled out	Ruled out	Ruled out

Table 8.2 Determining the status of the causal claim from causation in the study population and causation in the target population on the basis of evidence obtained on the target population. It is assumed here that similarity of mechanism has been established

		Causation in the study population				
		Established	Provisionally established	Arguable	Speculative	Ruled out
Causation from target studies	Established	Established	Established	Established	Established	Provisionally established
	Provisionally established	Established	Provisionally established	Provisionally established	Provisionally established	Arguable
	Arguable	Established	Provisionally established	Arguable	Arguable	Speculative
	Speculative	Established	Arguable	Speculative	Speculative	Ruled out
	Ruled out	Provisionally ruled out	Ruled out	Ruled out	Ruled out	Ruled out

There is a broad agreement between the approach presented here and that of IARC. As with the approach advocated here, IARC employs evidence of mechanisms to draw conclusions about causation at two places: to evaluate efficacy in humans on basis of evidence directly in humans and to ensure that causal claims in specific animal populations can be extrapolated to humans. For the first task, IARC employs the Hill indicators without assessing mechanistic studies in a systematic way. It is only in assessing external validity that IARC explicitly evaluates studies that investigate the details of the mechanism of action.

The approach presented here is more explicit with respect to where and what evidence of mechanisms should be used. Firstly, this book recommends explicitly evaluating mechanistic studies when evaluating evidence obtained directly in humans. After all, evaluating both whether there exists a mechanism and whether

there exists a correlation is necessary for evaluating the evidence obtained directly in humans (Sect. 7.1). The Hill indicators can only be seen as a first approximation to the comprehensive assessment of mechanistic evidence needed to establish efficacy in humans. What is more, these indicators tend to obfuscate, rather than clarify, distinctions between evidence pertinent to the correlational claim and evidence pertinent to the general mechanistic hypothesis (Chap. 6).

Secondly, this book separates the overall evaluation of causality and the evaluation of evidence directly obtained in humans. The overall evaluation is obtained by aggregating the evidence directly obtained in humans and the evidence in animals (Sect. 7.2). For instance, it might be that, initially, some causal claim is established in humans by considering studies that purely involve humans, but that, subsequently, studies of a variety of animal species that are mechanistically similar to humans rule out causation in those species. These further studies would surely cast enough doubt on causation in humans so that the causal claim can no longer be considered established. However, by identifying the overall evaluation with the evaluation of evidence directly obtained in humans when the evidence obtained on humans is sufficient (see the top row of the IARC table, Fig. 8.1), IARC assigns Group 1 in this case (the top right-hand corner of the IARC table). The procedure set out in this book would assign status *established* to the causal claim on basis of just the evidence directly obtained in humans, but it would assign overall status *provisionally established* on the basis of all the evidence, animal as well as human (see the top-right corner of Table 8.2). This classification is perhaps more appropriate.

8.2 Comparison to SYRINA

SYRINA is a framework that was put forward to evaluate the strength of evidence that a certain exposure is an endocrine disruptor (Vandenberg et al. 2016). This approach first evaluates the evidence for an association between chemical exposure and (adverse) effect. Second, this approach evaluates the evidence for an association between the chemical and endocrine disrupting activity. Third, the evidence for an association with an (adverse) effect and for an endocrine disrupting activity are combined to obtain an overall assessment of endocrine disruption.

SYRINA combines quality of evidence ratings from different streams of evidence in all three steps. As with our approach, the quality level of the causal claim is the minimum of the quality of the different evidence streams. Figure 8.2 gives the relevant SYRINA table for an association between chemical exposure and (adverse) effect.

The resulting initial rating can be upgraded by one level if there is high confidence in the evidence from *in silico* and *in vitro* studies.

In the next step, the endocrine disrupting activity of the exposure is evaluated by combining different evidence streams. This *in vivo* and *in vitro* evidence is combined. Figure 8.3 gives the relevant SYRINA table.

Fig. 8.2 SYRINA table for an association between chemical exposure and (adverse) effect

Human/Wildlife (observational)	High	Strong	Strong	Strong	Strong
	Medium	Moderate	Moderate	Moderate	Strong
	Low	Weak	Weak	Moderate	Strong
	Absent	No data	Weak	Moderate	Strong
	Absent	Low	Medium	High	
Experimental <i>in vivo</i>					

Fig. 8.3 SYRINA table for combining *in vivo* and *in vitro* evidence

<i>In vitro</i>	High	Strong	Strong	Strong	Strong
	Medium	Moderate	Moderate	Moderate	Strong
	Low	Weak	Weak	Moderate	Strong
	Absent	No data	Weak	Moderate	Strong
	Absent	Low	Medium	High	
Experimental <i>in vivo</i>					

Finally, the quality levels for the association with adverse health outcomes and for the endocrine activity are combined according to the table in Fig. 8.4.

In relatively unusual cases the resulting quality level can be upgraded or downgraded by considerations given to the plausibility of the link of disrupting endocrine disrupting activity and outcome.

Let us consider some points of comparison between SYRINA and the approach of this book. First, this book formulates explicit methods for evaluating evidence of mechanisms (Chap. 6). Second, for the evaluation of both endocrine activity and association with adverse health outcomes, SYRINA only combines two kinds of study. When evaluating the plausibility of an association with adverse outcomes, SYRINA combines results from experimental laboratory animals with evidence in humans or wildlife animals. According to the approach presented in this book, application of results from such associations in animals would need to be extrapolated with the help of evidence of mechanisms along the lines of Sect. 7.2. In addition, mechanistic considerations may be relevant when evaluating whether there is an association of the chemical with adverse health outcomes. After all, an observed correlation may be due to confounding. As with IARC, SYRINA makes use of the Hill indicators

Strength of Evidence: Endocrine Disrupting Activity	Strong	Probable EDC	Probable EDC	Known EDC	Known EDC
	Moderate	Possible EDC	Possible EDC	Probable EDC	Known EDC
	Weak	Not classifiable	Not classifiable	Possible EDC	Probable EDC
	No data	Not classifiable	Not classifiable	Possible EDC	Probable EDC
	No data	Weak	Moderate	Strength of Evidence: Association Between Exposure and Health Outcome	

Fig. 8.4 SYRINA table for combining the quality levels for the association and the endocrine activity

for evaluating each stream of evidence and does not explicitly distinguish between evidence of mechanisms and evidence of correlation. Hence, while this book agrees with SYRINA that many evidence streams should be considered when evaluating causal claims, we would emphasise the need for a more systematic integration of evidence of mechanisms and evidence of correlation along the lines of Chaps.6 and 7.

References

IARC (2015). *IARC monographs on the evaluation of carcinogenic risks to humans: Preamble*. Lyon: International Agency for Research on Cancer.

Hill, A. B. (1965). The environment and disease: Association or causation? *Proceedings of the Royal Society of Medicine*, 58, 295–300.

IARC (2009). *IARC monographs on the evaluation of carcinogenic risks to humans, vol. 100F: A review of human carcinogens. Chemical Agents and Related Occupations*. International Agency for Research on Cancer, Lyon.

Smith, M. T., Guyton, K. Z., Gibbons, C. F., Fritz, J. M., Portier, C. J., Rusyn, I., et al. (2016). Key characteristics of carcinogens as a basis for organizing data on mechanisms of carcinogenesis. *Environmental Health Perspectives*, 124, 713–21.

Vandenberg, L., Agerstrand, M., Beronius, A., Beausoleil, C., Bergman, A., Bero, L., et al. (2016). A proposed framework for the systematic review and integrated assessment (SYRINA) of endocrine disrupting chemicals. *Environmental Health*, 15(74), 1–19.

Wilde, M., & Parkkinen, V. -P. (2017). Extrapolation and the Russo–Williamson thesis. *Synthese*. <https://doi.org/10.1007/s11229-017-1573-y>.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

