A BAYESIAN REFLECTION ON THE MEANING OF EVIDENCE

ULLRIKA SAHLIN

LUND UNIVERSITY CENTRE OF ENVIRONMENTAL AND CLIMATE RESEARCH

Bayes@Lund 2016

ullrika.sahlin@cec.lu.se

Based on the body of evidence, we can see that it is absolutely correct to take a precautionary approach and ban these chemicals.

European Environment Agency

Topics | Data and maps | Indicators | Publications

You Explore by environmental topics ... are a huge risk – so ban is welcome, says EEA

Neonicotinoid pesticides are a huge risk – so ban is welcome, says EEA

Change language

Topics: Chemicals | Biodiversity | Agriculture

The European Commission has decided to ban three neonicotinoid insecticides. These chemicals can harm honeybees, according to a large body of scientific evidence, so the European Environment Agency (EEA) commends the precautionary decision to ban them.

> **Based on the body of evidence, we can see that it is absolutely correct to take a precautionary approach and ban these chemicals.**
>
> Jacqueline McGlade, EEA Executive Director

The three banned insecticides are clothianidin, imidacloprid and thiametoxam. A recent assessment from the European Food Safety Authority also found that there were "high acute risks" from the three insecticides.

In the recent EEA report 'Late Lessons from Early Warnings, volume II', published in January this year, the EEA considered the body of evidence surrounding imidacloprid from scientific studies, beekeepers and industry, concluding that the chemical should be withdrawn from the market given the evidence of harm and scale of the risk. The insecticides may directly affect a wide range of organisms, both on land and in water. In addition, honeybees and other insects perform vital pollination to crops

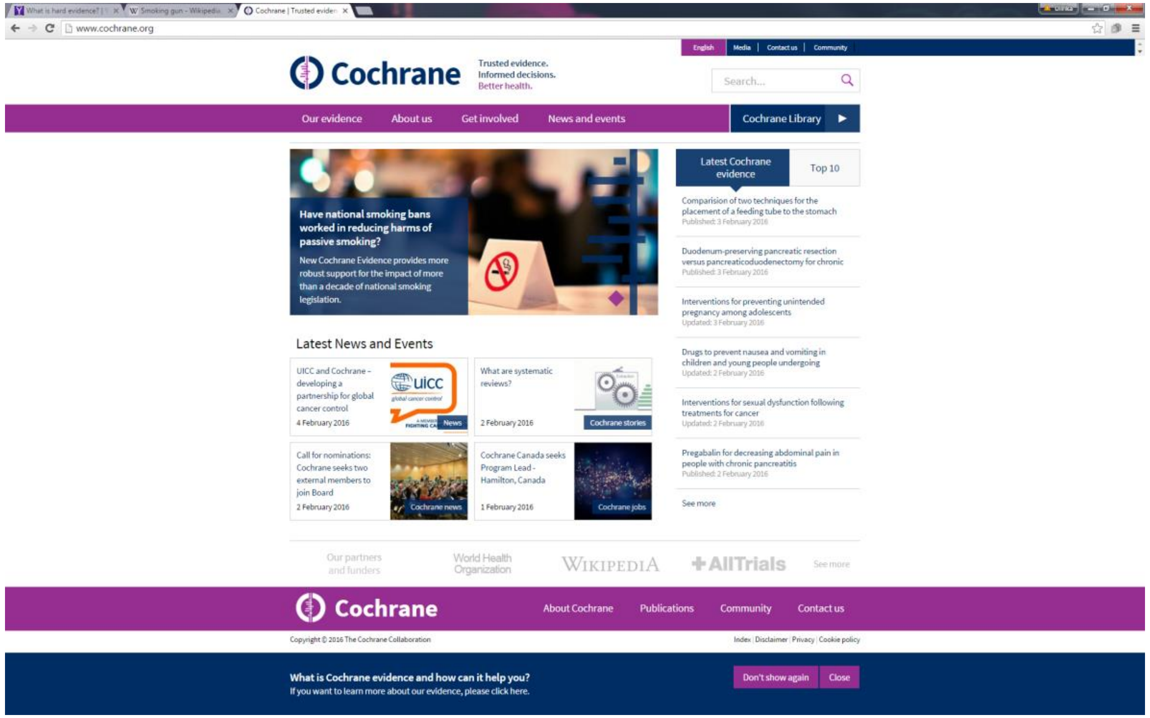
OUTLINE

- What is evidence
- Is there a ”Bayesian use” of evidence
- Is the world ready for a Bayesian approach to treat evidence and uncertainty
- Concluding remarks

WHAT IS EVIDENCE

- A thing or set of things helpful in forming a conclusion or judgment
- Ground for belief or disbelief
- Data on which to base proof or to establish truth or falsehood
- Something that makes plain or clear; an indication or sign
- *Law.* data presented to a court or jury in proof of the facts in issue and which may include the testimony of witnesses, records, documents, or objects.
- Evidence is research findings derived from the systematic collection of data through observation and experiment and the formulation of questions and testing of hypotheses
- Quantify evidence in favor of the null-hypothesis
- Synonyms: information, …, proof

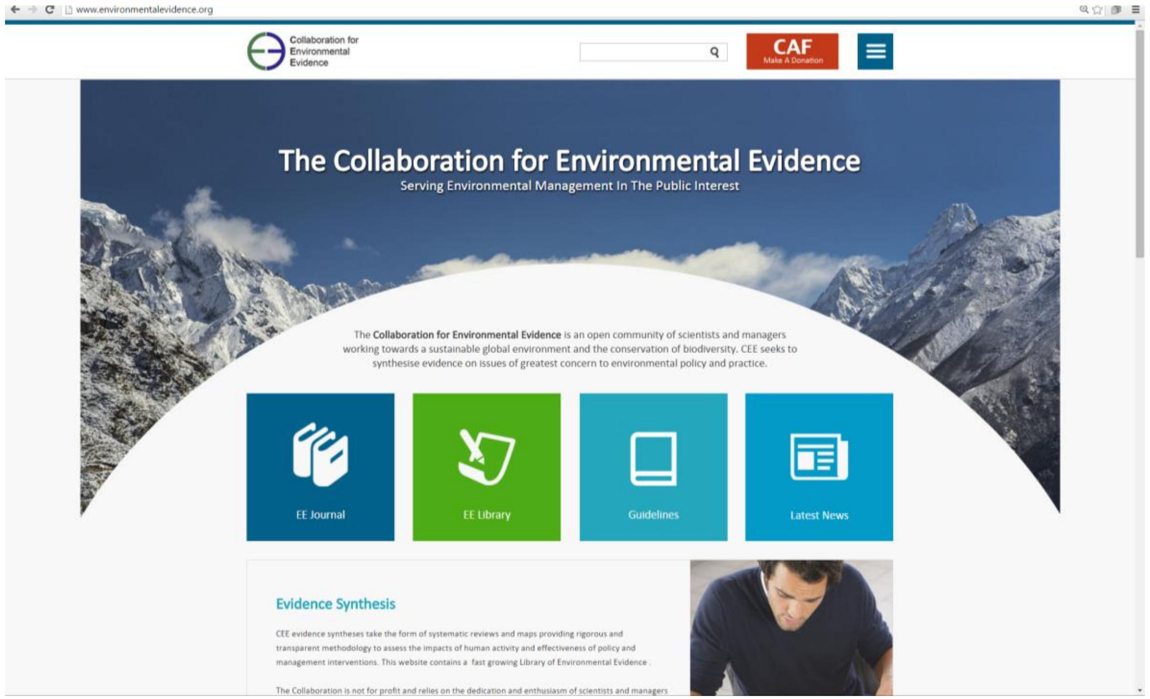
MEDICINE



GRADE

Underlying methodology	Quality rating
Randomized trials; or double-upgraded observational studies.	High
Downgraded randomized trials; or upgraded observational studies.	Moderate
Double-downgraded randomized trials; or observational studies.	Low
Triple-downgraded randomized trials; or downgraded observational studies; or case series/case reports.	Very low

ENVIRONMENTAL MANAGEMENT DECISIONS



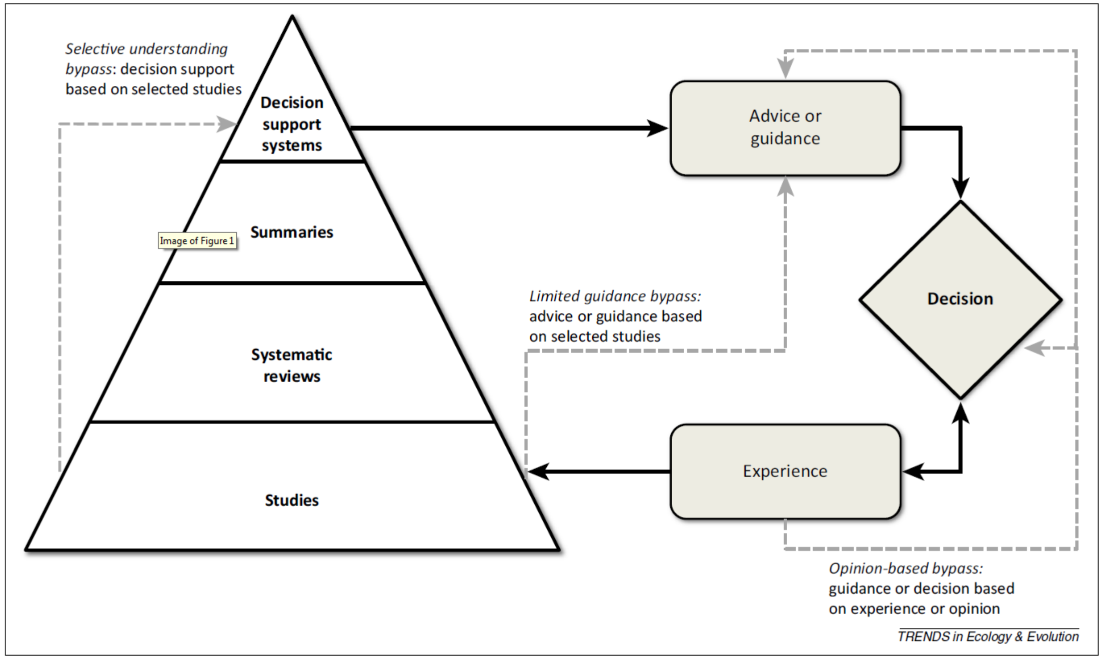


Figure 1. A schematic showing how scientific information could feed into environmental decisions. The triangle on the left is a simplification of the '4S' or '5S' hierarchy proposed by Haynes [22,24] in which summaries integrate evidence from studies and systematic reviews, and are used as the basis for information flowing into decision support systems. In this scheme, environmental decisions are based on the best-available evidence, combined with the expertise and local knowledge of the practitioner or policymaker (described by the 'Experience' box). Broken lines illustrate bypass routes currently taken to inform environmental decisions (see main text for details).

Dicks, L., Walsh, J. and Sutherland, W. (2014). Organising evidence for environmental management decisions: a "4S" hierarchy. Trends in Ecology & Evolution 29:607-612.

META-ANALYSIS

Charnley

Stanmore

Table IV. Summary of evidence on revision hazards for Charnley and Stanmore prostheses: hazard ratios < 1 are in favour of Stanmore.

Source	Charnley		Stanmore		Estimated hazard ratio	
	Number of patients	Revision rate	Number of patients	Revision rate	(HR)	(95% int.)
					Fixed-effects model	
Registry	28 525	5.9%	865	3.2%	0.55	(0.37–0.77)
RCT	200	3.5%	213	4.0%	1.34	(0.45–3.46)
Case series	208	16.0%	982	7.0%	0.44	(0.28–0.66)
					Common-effect model	
					0.52	(0.39–0.67)
Quality weights [registry, RCT, case series]					*Random-effects model*	
				[1, 1, 1]	0.54	(0.37–0.78)
				[0.5, 1, 0.2]	0.61	(0.36–0.98)
				[0.1, 1, 0.05]	0.82	(0.36–1.67)

Spiegelhalter and Best (2003). Bayesian approaches to mulitple sources of evidence and uncertainty in complex cost-effectiveness modelling. *Statist. Med.*

”BAYESIAN USE” OF EVIDENCE – BAYESIAN BELIEF NETWORKS

- *Hard* evidence (instantiation) for a node X in a BBN is evidence that the state of X is definitely a particular value.
- *Soft* evidence for a node X in a BBN is any evidence that enables us to update the prior probability values for the states of X.
- “Soft evidence does still have value, but only in the absence of hard evidence. If there is no supporting hard evidence, then the best available forms of soft evidence should be used in the meantime.”

”BAYESIAN USE” OF EVIDENCE – EXPERT INFORMED PRIORS

- Can expert knowledge be evidence?

”BAYESIAN USE” OF EVIDENCE - BAYESIAN META-ANALYSIS

Table IV. Summary of evidence on revision hazards for Charnley and Stanmore prostheses: hazard ratios < 1 are in favour of Stanmore.

	Charnley		Stanmore		Estimated hazard ratio	
Source	Number of patients	Revision rate	Number of patients	Revision rate	(HR)	(95% int.)
					Fixed-effects model	
Registry	28 525	5.9%	865	3.2%	0.55	(0.37–0.77)
RCT	200	3.5%	213	4.0%	1.34	(0.45–3.46)
Case series	208	16.0%	982	7.0%	0.44	(0.28–0.66)
					Common-effect model	
					0.52	(0.39–0.67)
Quality weights [registry, RCT, case series]					*Random-effects model*	
				[1, 1, 1]	0.54	(0.37–0.78)
				[0.5, 1, 0.2]	0.61	(0.36–0.98)
				[0.1, 1, 0.05]	0.82	(0.36–1.67)

Spiegelhalter and Best (2003). Bayesian approaches to mulitple sources of evidence and uncertainty in complex cost-effectiveness modelling. *Statist. Med.*

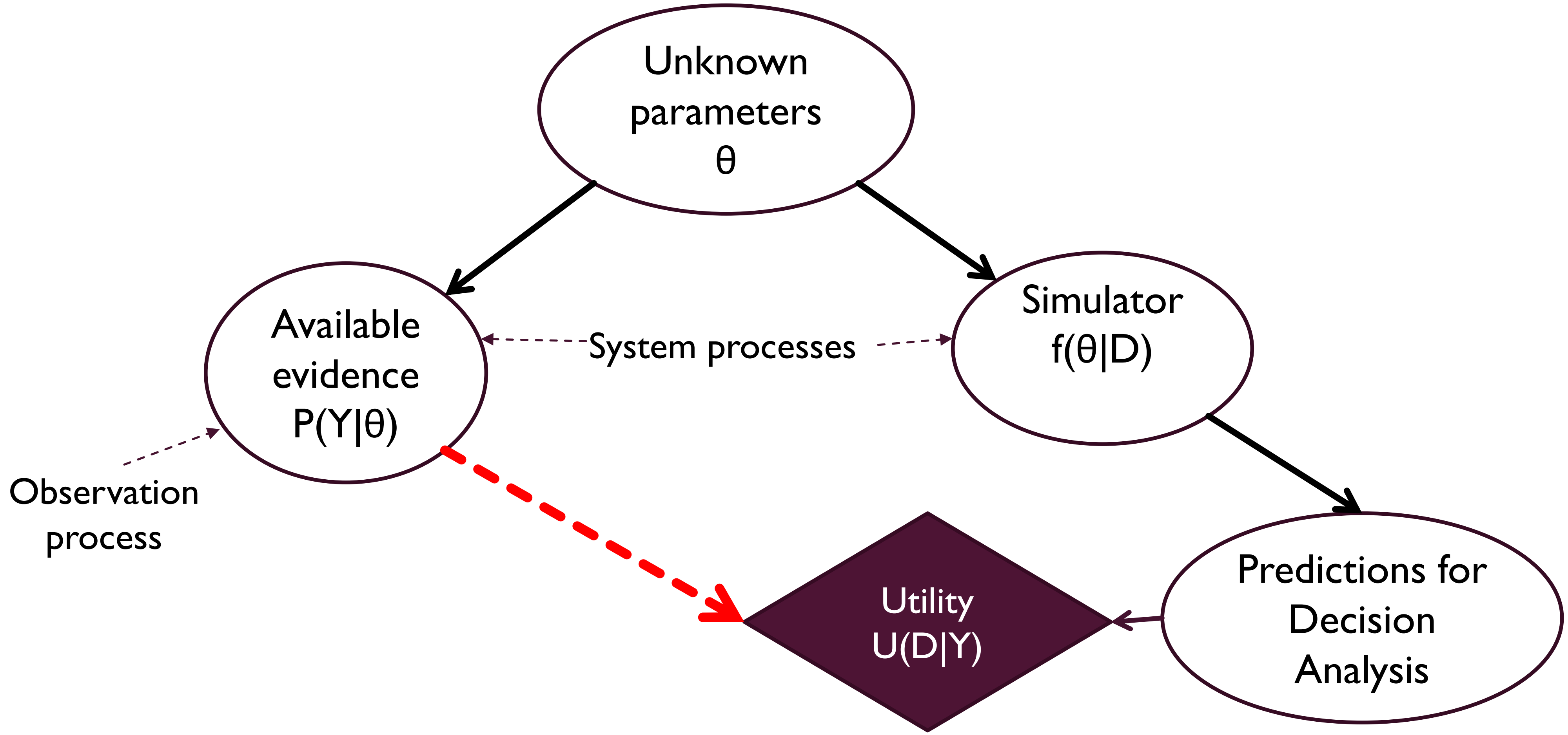
"BAYESIAN USE" OF EVIDENCE – BAYESIAN HIERARCHICAL MODELLING

- BHM to take into account differences between sources of evidence
- System process
- Observation process
- Extrapolation process
- …

Table IV. Summary of evidence on revision hazards for Charnley and Stanmore prostheses: hazard ratios < 1 are in favour of Stanmore.

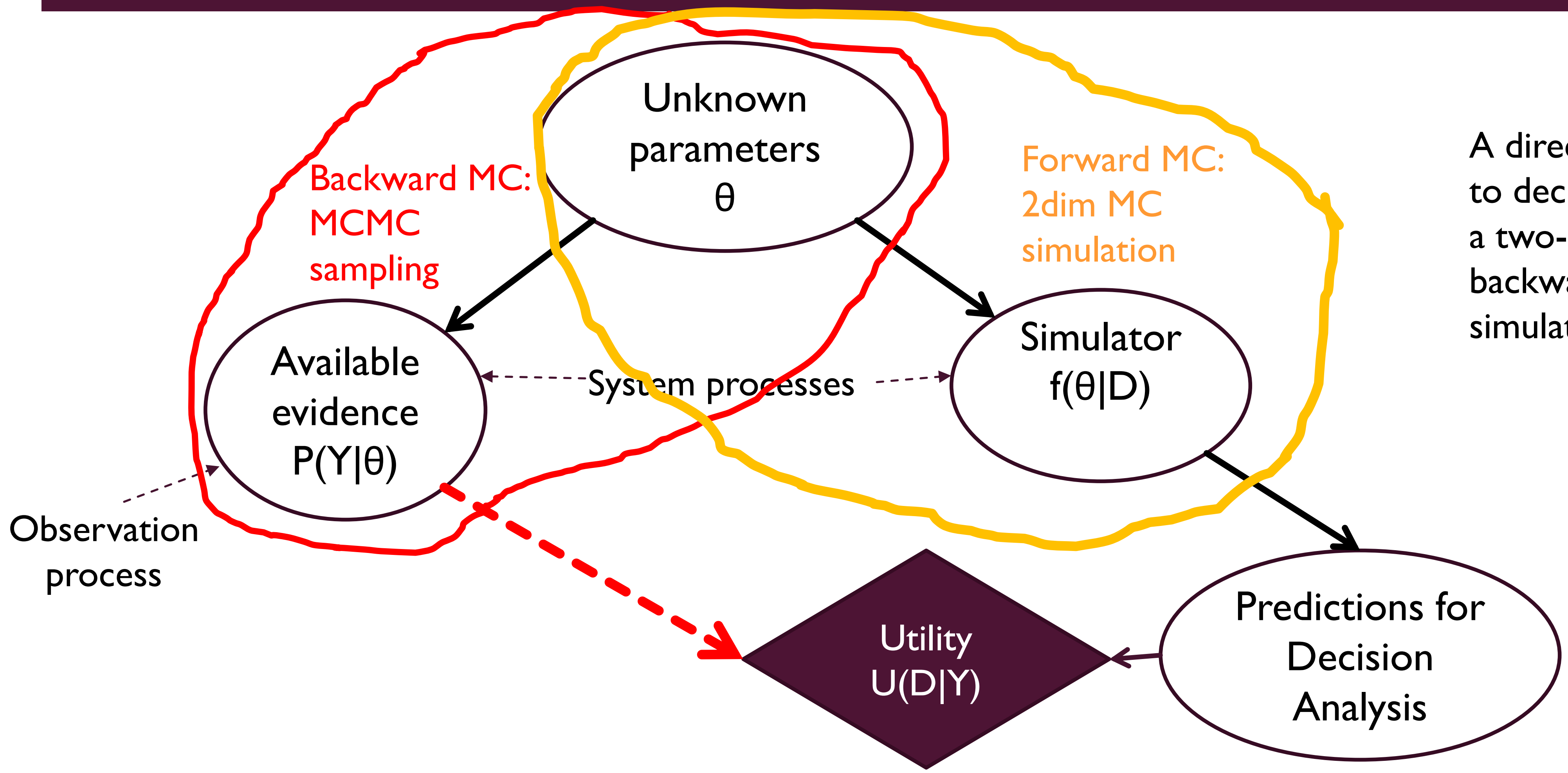
Source	Charnley: Number of patients	Charnley: Revision rate	Stanmore: Number of patients	Stanmore: Revision rate	Estimated hazard ratio (HR)	(95% int.)
					Fixed-effects model	
Registry	28 525	5.9%	865	3.2%	0.55	(0.37–0.77)
RCT	200	3.5%	213	4.0%	1.34	(0.45–3.46)
Case series	208	16.0%	982	7.0%	0.44	(0.28–0.66)
					Common-effect model	
					0.52	(0.39–0.67)
Quality weights [registry, RCT, case series]					*Random-effects model*	
				[1, 1, 1]	0.54	(0.37–0.78)
				[0.5, 1, 0.2]	0.61	(0.36–0.98)
				[0.1, 1, 0.05]	0.82	(0.36–1.67)

"BAYESIAN USE" OF EVIDENCE – BAYESIAN EVIDENCE SYNTHESIS



Spiegelhalter and Best (2003). Bayesian approaches to mulitple sources of evidence and uncertainty in complex cost-effectiveness modelling. *Statist. Med.*

”BAYESIAN USE” OF EVIDENCE – BAYESIAN EVIDENCE SYNTHESIS



A direct link from data to decision instead of a a two-stage approach: backward and forward simulation

Spiegelhalter and Best (2003). Bayesian approaches to mulitple sources of evidence and uncertainty in complex cost-effectiveness modelling. *Statist. Med.*

IS THE WORLD READY FOR A BAYESIAN TREATMENT OF EVIDENCE AND UNCERTAINTY

First, the world must be aware of that there is uncertainty

A TAXONOMY AND TREATMENT OF UNCERTAINTY

TABLE 1. The various sources of epistemic and linguistic uncertainty with their most appropriate general treatments (refer to relevant section for references related to the suggested treatment).

Source of uncertainty	General treatments
Epistemic uncertainty	
Measurement error	statistical techniques; intervals
Systematic error	recognize and remove bias
Natural variation	probability distributions; intervals
Inherent randomness	probability distributions
Model uncertainty	validation; revision of theory based on observation; analytic error estimation (for meta-models)
Subjective judgment	degrees of belief; imprecise probabilities
Linguistic uncertainty	
Numerical vagueness	sharp delineation; supervaluations; fuzzy sets; intuitionistic, three-valued, fuzzy, paraconsistent and modal logics; rough sets
Nonnumerical vagueness	construct multidimensional measures then treat as for numerical vagueness
Context dependence	specify context
Ambiguity	clarify meaning
Indeterminacy in theoretical terms	make decision about future usage of term when need arises
Underspecificity	provide narrowest bounds; specify all available data

Regan et al.

IS THE WORLD READY FOR A BAYESIAN TREATMENT OF EVIDENCE AND UNCERTAINTY

Second, adapt scientific method to use principles to quantify uncertainty when that is the objective

Stated 1990! Is it done today?

The Concept of Probability in Safety Assessments of Technological Systems

George Apostolakis

Safety assessments of technological systems, such as nuclear power plants, chemical process facilities, and hazardous waste repositories, require the investigation of the occurrence and consequences of rare events. The subjectivistic (Bayesian) theory of probability is the appropriate framework within which expert opinions, which are essential to the quantification process, can be combined with experimental results and statistical observations to produce quantitative measures of the risks from these systems. A distinction is made between uncertainties in physical models and state-of-knowledge uncertainties about the parameters and assumptions of these models. The proper role of past and future relative frequencies and several issues associated with the elicitation and use of expert opinions are discussed.

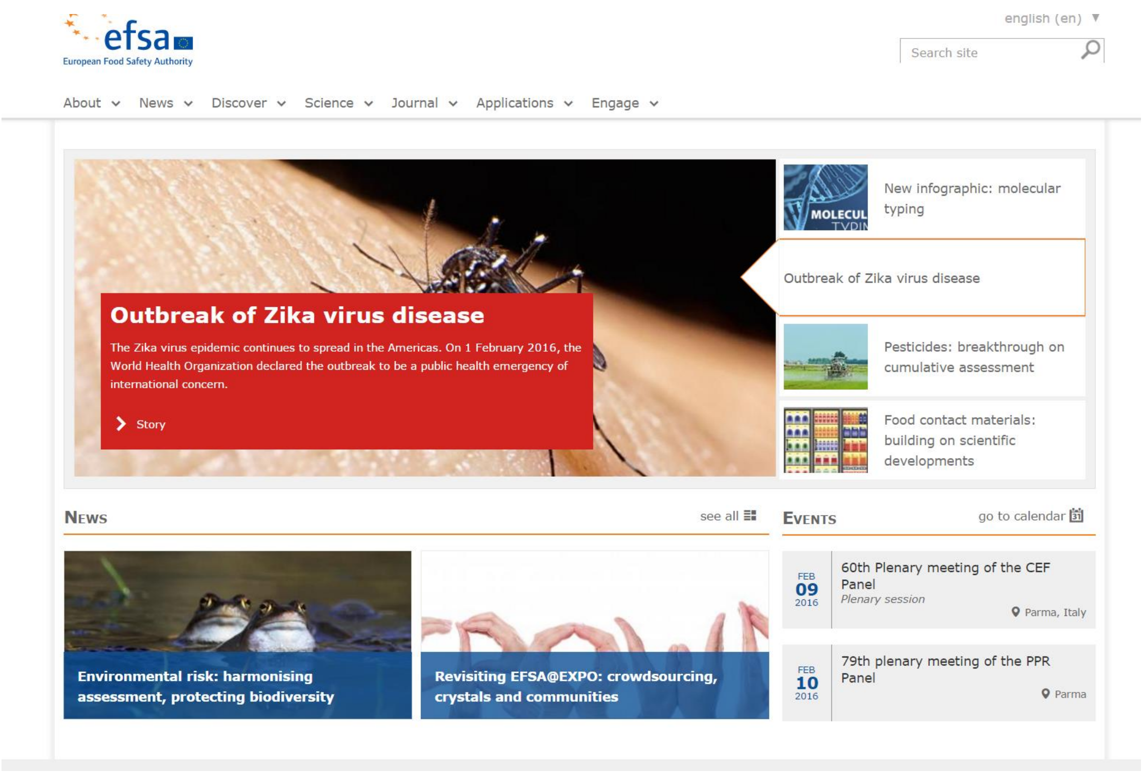
Probabilistic risk assessment (PRA) or probabilistic

among alternatives). The first element (problem-structuri the foundation upon which one performs further analysis b ing models for the physical world and developing alternative of action. The second element requires the introduction o bilities and their calculus. The preferences (third elem expressed in terms of utilities and, finally, the decision criteri maximization of the expected utility (fourth element). A per follows this procedure in decision-making and whose prob comply with the theory of probability is a coherent decisio (*1–3*).

For major societal decisions that involve many decision (or, more accurately, many stakeholders), formal decision breaks down. Because this theory guarantees coherence probability assignments and preferences of a single decision two decision makers may be individually coherent and still b to agree and reach the same decision. In these situations the elements of the decision problem, that is, the quantific preferences and the maximization of expected utilities, are by ad hoc decision-making criteria that are widely debat ultimately, imposed by the regulatory authority. The p

IS THE WORLD READY FOR A BAYESIAN TREATMENT OF EVIDENCE AND UNCERTAINTY

- European Food Safety Authority – Environmental and Health risk assessments for the EU
- Guidance on Uncertainty in EFSA Scientific Assessment (draft Feb 2016)
- To meet the general requirement for transparency, all EFSA scientific assessments must include consideration of uncertainties



EFSA KEY CONCEPTS FOR UNCERTAINTY ANALYSIS

- *Uncertainty is personal and temporal.* The task of uncertainty analysis is to express the uncertainty of the assessors, at the time they conduct the assessment: there is no single ”true” uncertainty

Sounds very Bayesian to me

- *Evidence, agreement, confidence and conservatism* are related but distinct concepts. Measures of evidence and agreement may be useful in assessing uncertainty but are not sufficient alone.

What would be the relation between evidence and uncertainty in a Bayesian perspective

EFSA KEY CONCEPTS FOR UNCERTAINTY ANALYSIS

- *Probability* is the preferred measure for expressing uncertainty, as it quantifies the relative likelihood of alternative outcomes, which is what decision-makers need to know
- All well-defined uncertainties can be quantified using subjective probability

Why is takling about probabilty so complicated?
So is the subjective probability a different probabilty than the previous?
Did they mean that the previous is a relative frequency? Sometimes it is, sometimes not. If not, it is a subjective probability as well.

- Uncertainty analysis should begin early in the assessment process and not be left to end

Yes!

QUANTITATIVE METHODS REVIEWED BY EFSA – ”WELL KNOWN”

- Quantitative uncertainty tables
- Deterministic calculations with conservative assumptions
- Sensitivity analysis
- Interval analysis
- Expert knowledge elicitation
- Monte Carlo simulation – taking random samples from probability distributions representing uncertainty and/or variability

- *Confidence intervals & the Bootstrap* - quantifying uncertainty about parameters in a statistical model of variability on the basis of data
- *Bayesian inference* – quantifying uncertainty about parameters in a statistical model of variability on the basis of data and expert judgement about the values of the parameters

The Bayesian inference is taking into account!

QUANTITATIVE METHODS REVIEWED BY EFSA – ”ODD”

- Probability bound analysis - a general method for combining limited probablity specifications about inputs in order to make a limited probabiltiy specification about the ouput of a risk calculation.
- Other quantitative methods - uncertainty expressed in terms of
  - Possibilities
  - Imprecise probabilities
  - Bayesian modelling

Bayesian modelling - the last method mentioned
But Bayesian modelling is not an expression of uncertainty – it is a way to quantify uncertainty by probability

IS THE WORLD READY FOR A BAYESIAN TREATMENT OF EVIDENCE AND UNCERTAINTY

What does Spiegelhalter say?

RECOMMENDATIONS TO FACE DEEPER UNCERTAINTIES IN MODELLING FOR EVIDENCE AND DECISIONS

1. Use quantitative models with aleatory and epistemic uncertainty expressed as Bayesian probability distributions
2. Conduct sensitivity analysis to alternative model forms and assess evidential support for alterantive structures, without putting probabilities to models
3. Provide a list of known model limitations and a judgement of their qualitative or quantitative influence and ensuring there has been a fully imaginative consideration of possible futures
4. Provide a qualitative expression of confidence, or lack of it, in any analysis based on the quality of the underlying evidence, possibly expressed using an adpated GRADE scale or the IPCC guidance

Spiegelhalter and Riesch (2011). Don't know, can't know: embracing deeper uncertainties when analysing risks. Phil. Trans. R. Soc. A

Is the world ready for a Bayesian treatment of evidence and uncertainty

RECOMMENDATIONS TO FACE DEEPER UNCERTAINTIES IN MODELLING FOR EVIDENCE AND DECISIONS

5. In situations of low confidence, use deliberately imprecise expressions of uncertainty about quantities, such as their orders-of-magnitude, whether they are positive or negative, or even refuse to give any judgement at all; the IPCC guidance suggests a calibrated scale for these expressions
6. When exploring possible actions, look for robustness to error, reslience to the unforeseen, and potential for adaptivity in the face of the unexpected
7. Seek transparency and ease of interrogation of any model, with clear expression of the provenance of assumptions
8. Communicate the estimates with humility, communicate the uncertainty with confidence.

Spiegelhalter and Riesch (2011). Don't know, can't know: embracing deeper uncertainties when analysing risks. Phil. Trans. R. Soc. A

CONCLUDING REMARKS

A Bayesian perspective allow us to:

- Consider quality in evidence evaluted in different ways, including expert judgement
- Quantify uncertainty in evidence taking into account differences in quality
- Take into account different observations processes
- View ”more or less complex model based” predictions as evidence (perhaps at a different level e.g. soft versus hard evidence)
- Evalute robustness in evidence by combining BES with sensitivity analysis or use generalized BES