"E-Synthesis: a Bayesian framework for probabilistic causal assessment"

Abstract

The problem of collecting, analyzing and evaluating evidence for causal assessment is a central one in science-based policy. In particular, due to the ongoing decision-making process in policy-making, and the radical uncertainty affecting it, *probabilistic tools of causal assessment* are essential to allow flexible yet justified and robust inferential procedures.
The ERC-funded PhilPharm project focused on probabilistic causal assessment in pharmacovigilance. This constituted a formidable laboratory for the integration of causal (inference) theories, statistical paradigms, scientific methodology and epistemology.
We aim to develop E-Synthesis in further directions both "horizontally" (by applying it to other fields in policy making and more generally in science-based decision-making) and "vertically" (by further investigating its foundational implications and mathematical modelling). For this, we aim to gather scholarly expertise in probability and probabilistic reasoning, formal epistemology, rational choice theory, philosophy of science and scientific methodology, epistemology, sociology of science, science policy. We also aim to develop an open-source application integrating expert systems approaches of diverse kinds (Bayesian expert systems, AI, Machine Learning) in collaboration with possible end-users (e.g. Governmental Agencies, Policy-makers, scientific experts).
The project will aim to:

1) further investigate the foundational bases underpinning *E-Synthesis* (epistemic dynamics, probability kinematics, justification of causal inference, statistical data analysis, methodological issues and exogenous factors impacting on evidence quality). In particular, strategic dimensions in evidence sampling, analysis, disclosure and communication will be in focus (see Osimani, 2024);
2) further develop the *E-Synthesis* framework and implement it into a prototype software for decision support. As an immediate goal, we privilege the development of a decision aid for pharmacosuveillance, however, any other relevant field of application is considered (especially environmental policy).

E-Synthesis for pharmacosurveillance

Pharmaceutical safety is a central one in health-care practice. Adverse Drug Reactions (ADRs) are responsible for a heavy economic and social burden (Edwards and Lundkvist 2017); moreover, they constitute an extremely vulnerable point for the health system and a key ethical problem for decisions concerning pharmaceutical products — which stood as a top-priority also in the current Covid-19 pandemic (Chandler et al. 2020).
The European Parliament and the European Council regulation for pharmacovigilance practice (Directive 2010/84/EU; Regulation (EU) No 1235/2010) have been putting a special emphasis on joint efforts for what can be considered an information-based (rather than power-based) approach to pharmaceutical risk assessment. The related guidelines encourage the integration of information coming from different sources of safety signals (spontaneous case reports, data mining, pharmacoepidemiological studies, post-marketing trials, drug utilization studies, non-clinical studies, late-breaking information, see also Herxheimer 2012). In the U.S. the Congress approved Pub.L. 114 - 255 (21st Century Cures Act) in 2016. This specifies conditions that allow companies to provide "data summaries" and "real world evidence" such as observational studies, insurance claims data, patient input, and anecdotal data rather than exclusively relying on random clinical trials (RCTs) for the purpose of drug approval and evaluation. Yet, the methodological bases for implementing such policies are shaky, in that causal assessment of ADRs is still parasitic on the (statistical) methods developed to test drug efficacy (see also Senn 2007), which are more focused on hypothesis testing ("estimating the effects of causes"), rather than generation ("discovering causes of effects"), and therefore fall short of using sporadic, fragile and heterogeneous evidence efficiently.
The ERC project PhilPharm "Philosophy of Pharmacology: Safety, Statistical Standards, and Evidence

Amalgamation" GA (639276) has addressed this problem by developing a theoretical framework for probabilistic confirmation of causal hypotheses on the basis of all the available evidence on safety for medical treatments: "*E-Synthesis*" (Landes et al. 2018; Abdin et al. 2019; De Pretis and Osimani 2019; De Pretis et al. 2019; De Pretis et al. 2020). Here a recent interview on the project appeared on the CORDIS website.

E-Synthesis is to date the unique attempt to systematically synthetize *heterogenous* evidence in pharmacology. The framework aims to probabilistically measure the hypothesis of a causal connection between a given pharmaceutical treatment and a possible (side-)effect, in a given population for a given causal model (Landes et al. 2018; De Pretis et al. 2019). This is especially important in pharmacosurveillance because, whereas evaluating *intended* (beneficial) effects of interventions may rely on data pooling (meta-analyses), and systematic reviews of homogeneous studies, risk detection and discovery is characterized by 1) relative scarcity, 2) heterogeneity, and 3) "fragility" of data concerning the safety of health technologies.

In (Landes et al. 2018; De Pretis et al. 2019), we developed a theoretical framework where any kind of safety signal can contribute to the confirmation or disconfirmation of hypothesized causal associations between drug and adverse reactions, along possibly different lines of evidence (e.g. observational vs. experimental, or population level vs. clinical or biochemical level), exploiting their *joint* contribution to causal inference.

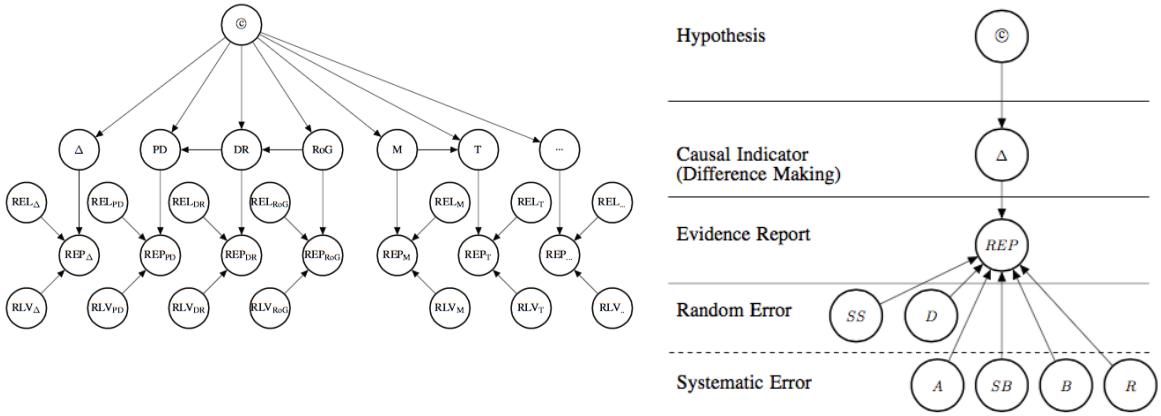

Fig. 1. The *E-Synthesis* framework. Left: the Bayesian network represents the hypothesis of causal association between a drug and a suspected side effect, denoted by ©, as an ancestral node, its direct children are abstract indicators of causation — such as "Difference Making" (**Δ**); Probabilistic Dependence (PD); Dose Response (DR); Rate of Growth (RoG); underpinning mechanisms (M), time precedence (T). Causal indicators are themselves "parents" of report nodes (REP_i); that is, study results or any other data. REL and RLV nodes refer to Reliability and Relevance (i.e. internal and external validity, broadly speaking). Once evidence comes in, probabilities propagates through the network and change the probability of causation ©. **Right:** a "section" of the Bayesian network where the REL and RLV nodes are "exploded" into various indicators of internal and external validity (SS: sample size; D: study duration; A: adjustment/stratification; SB: sponsorship bias; R: randomization). Source: Landes, Osimani, Poellinger 2018, and De Pretis, Landes Osimani, 2019.

We report progress towards concrete implementation in (Abdin et al. 2019; De Pretis et al. 2019; De Pretis and Osimani 2019; De Pretis et al. 2020). There, we begin to determine the decision weights of different types of evidence in concrete applications (Abdin et al. 2019; De Pretis et al. 2019; De Pretis et al. 2020) and we focused on dose-response evidence, providing a new computational model to analyze it for pharmacovigilance purposes (De Pretis and Osimani 2019).

The present project will be devoted to:

3) further investigate the foundational bases underpinning the framework (epistemic dynamics, probability kinematics, justification of causal inference, statistical data analysis, methodological issues and exogenous factors impacting on evidence quality). In particular, strategic dimensions in evidence sampling, analysis, disclosure and communication will be in focus (see Osimani, 2024);
4) further develop the *E-Synthesis* framework and implement it into a prototype software for decision support. As an immediate goal, we privilege the development of a decision aid for pharmacosuveillance, however, any other relevant field of application is considered (especially environmental policy).

Pharmacosurveillance Open Souce Software

E-Synthesis has been presented at numerous workshops and conferences and its development has been made possible by the collaborative work of several expertise both in the academia (epistemologists, methodologists, mathematicians) and in the third sector (governmental agencies). In particular, we could profit from deep interviews and discussions with various representatives of European Drug Agencies and Health Scientists in several occasions: during the focus group meetings of the ERC PhilPharm project at the MuST conference in Munich; as well as at a workshop organized by the PhilPharm team in January 2017: "Drug Safety, Probabilistic Causal Assessment, and Evidence Synthesis" (Munich), the Joint Conference on Biometrics & Biopharmaceutical Statistics 2017, Vienna and the roundtable "Evidence in Biomedical, Statistical and Legal Sciences" 2018, Ancona. *E-Synthesis* has been also presented at the Issues in Medical Epistemology Conference 2017 in Cologne. Furthermore, Barbara Osimani, PI of the project, has been invited to present it at the EMA in London (23rd November 2016), at the French Drug Agency (28th June 2017 and 6th July 2018), at the German Drug Agency (11th July 2017) and the Austrian Drug Agency (30th August 2017, 2nd and 3rd July 2018), WHO affiliated Uppsala Monitoring Center (23rd and 24th May 2018) and the Italian Drug Agency (30th September 2019). Follow-up meetings between the team and the French Drug Agency took placed in Paris on 22nd May 2019 and 26th June 2019.
Also the IT-industry and the Engineering Faculty at Univpm have been actively involved (Emanuele Frontoni, Adriano Mancini, as well as Aldo Dragoni and his research group).
The pharmacosurveillance implementation of E-Synthesis into a software package has been focusing on the line of evidence concerning the dose-response relationship[1] between drug and effect. In recent years an extensive literature has been developed regarding dose-finding assessment, a task that is usually targeted in Phase II of clinical research. A frequentist-grounded algorithm named Multiple Comparison Procedures and Modeling (MCP-Mod) has shown promising features in solving this task and has been object of several developments (Bretz et al. 2005; Pinheiro et al. 2014; Schorning et al. 2016; Bornkamp et al. 2017), being eventually qualified by the European Medicines Agency (EMA) in 2014 (see: EMA 2014).
A Bayesian approach for dose-finding has been also proposed in several works (Shao and Shapiro 2018; Liu and Johnson 2016; Takeda and Morita 2018; Toumazi et al. 2018; Mu and Xu 2017). In particular, in (Shao and Shapiro 2018) we found a suitable modeling for analyzing dose-response evidence and adapted it to the *E-Synthesis* Bayesian framework in De Pretis and Osimani 2019.

We propose the codification of a prototype software for the estimation, validation and update of the probability that a drug and a suspected harm are truly related by a non-spurious causal relationship. This prototype software aims to enable: 1) ingestion of data (e.g. dose-response reports) also considering population variables as age, sex, etc., 2) visualization of clustered data and 3) model estimation/validation/update. In case new reports become available (and then ingested) existing model(s) will be updated and the hypothesis "Drug D causes harm E in population U and model M" can be probabilistically quantified.

Economic and Societal Benefits

The endemic uncertainty affecting scientific knowledge and the urgent need of flexible instrument for policy-making require dynamic tools of hypothesis update, which are capable of tracking different types of uncertainty at various levels, and allow experts and decision-makers to consider the joint evidential support of any type of data into their evaluations. As a probabilistic tool for scientific inference, E-Synthesis perfectly meets these criteria.

Regarding the specific implementation of E-Synthesis for pharmacovigilance, one cannot enough underline its relevance and urgency. ADRs represent an open problem to drug manufacturers, are a serious risk to patients, and constitute a major ethical problem in licensing decisions of pharmaceutical products. In the US, 3 to 7% of all hospitalizations are due to adverse drug reactions, while ADRs occur during 10 to 20% of hospitalizations; about 10 to 20% of these ADRs are severe and can result in death. (Weiss et al. 2018). About 197,000 people in the EU die each year as a result of ADRs (Bouvy et al. 2015). Their clinical and economic burden is massive costing around

[1] "Dose–response relationship", is standardly defined as the magnitude of the response of an organism, as a function of exposure (or doses) to a stimulus or stressor (usually a chemical) after a certain exposure time (Crump et al. 1976). In *E-Synthesis* we aim to model any evidence of dose-response at various levels of the organism (molecular, clinical, epidemiological). Hence, evidence of dose-response may also come from an observational study, whose arms instantiate different levels of exposure, in terms e.g. of daily dosage, or, more generally, frequency of use of a given drug.

USD 30.1 billion only in the US (Sultana et al. 2013). The problem is widespread and affect both developed (Walters et al. 2016; Rolfes et al. 2016; Gyllensten et al. 2014) and developing countries (Fasipe et al. 2018) as well. Pharmacovigilance business has a grounded potential and accrues for about USD 12 billion globally (Fortune Business Insights 2020). The ERC project PhilPharm (GA 639276) has responded to the methodological challenges faced in this field by developing the foundational underpinnings of causal assessment in pharmacology, and laid the ground for the development of innovative tools (software products) for decision makers in drug licensing agencies. A successful implementation to better predict ADRs and improve the risk management process (by minimizing both false positives and false negatives) will benefit the patients, society, and the pharmaceutical industry as well.

The impact of this software would directly help the post-marketing safety assessment of medications by improving the process of data-analysis and evidence evaluation both for regulatory agencies and pharmaceutical companies, thereby making pharmacosurveillance more efficient and transparent. By making all dimensions and types of evidence explicit and synthesizing them into a unique framework, *E-Synthesis* also allows decisions to be more structured and therefore the related software would considerably ease sensitivity analyses based on changing evidence weights and modulators. The software would have an indirect impact by improving safety of patients and offering other societal benefits, like fewer hospitalizations due to ADRs and a better care of chronic patients that are subjected to continuous drug administration.

Competitive Analysis

Hazard measurement is generally considered to be complex and fraught with uncertainties (Osimani 2010). Numerous guidelines for the estimation of causal relationships between hazards and technologies have been developed. However, none of them undertakes the task to *probabilistically estimate* such causal relationships based on *all available (meta-)data*. Ad-hoc guidelines for specific products or chemicals are available based on heuristics, clinical judgment, and statistical canons unsystematically applied (see e.g. Shanks et al. 2009). Some of the tools extrapolate animal data to estimate effects on humans (ECETOC 2009).
The EMA-HMA guidelines on pharmacosurveillance (Module VII-X) encourage the integration of information coming from different sources of safety signals. However, despite these guidelines, there is still a lack of concrete implementations. A comprehensive tool, which integrates heterogeneous items of evidence in a systematic, transparent and traceable way is still missing. Indeed, although guidelines for harm assessment and semi-formal tools that permit the incorporation of some safety signals do exist, pharmacovigilance experts emphasized (in their meetings with us) the lack of formal quantitative tools aiding the assessment of drug induced harms on the basis of all available evidence. Within the timeframe of this PoC, we aim to develop the dose-response component of the software.
The overarching framework is Bayesian because the Bayesian theory of statistical inference may provide a better match to the dynamic and complex nature of pharmacosurveillance (see for instance Rodrigues et al. 2018), a context in which knowledge accrues across multiple studies and observations, and the hypothesis is generated and confirmed by the joint contribution of heterogeneous items of evidence (Osimani and Mignini 2015, Lesaffre 2012; Berry et al. 2010; Spiegelhalter 2006).