

PRACTICAL REASONING FOR INDIVIDUAL RISK-BENEFIT ASSESSMENT OF FOODS AND DRUGS

Sabina Costache

*Department of Computer Science and Information Technology, Dunarea de Jos
University of Galati*

Abstract: The paper introduces the new challenges of benefit risk assessment of foods and drugs in modern personalized medicine. A new modeling of benefit -risk assessment with valued argumentation frameworks is presented and exemplified, and further research directions are suggested, as nutrigenomics will discover new bio-mechanisms relevant to nutrition and toxicity.

Keywords: argumentation framework, non-monotonic reasoning, risk-benefit assessment, nutrigenomics.

1.INTRODUCTION

Risk –benefit assessment of foods and drugs is a new, challenging domain that has the aim to integrate information from other different research traditions (epidemiology, toxicology, nutrition, risk-assessment, nutrigenomics), in order to answer the question if certain beneficial foods/micronutrients are worth the risk of being included in a recommended diet (considering the adverse effects they simultaneously produce -due to either unavoidable contaminants or even intrinsic components). The root of this domain resides in the idea that “good food and nutrition can improve health and that some risk may be acceptable if benefit is expected to outweigh it” –Tijhuis (2012).

Risk assessment is addressed by toxicology and epidemiology. The traditional risk-assessment procedure consists of -Tijhuis (2012) :

- Hazard identification (what effect?)
- Hazard characterization (at what dose? How?)
- Exposure assessment (how much is taken in)
- Risk characterization (what is the probability and severity of the effect)

In toxicology, the characterization of dose-response relationships for toxic non-carcinogenic substances considers the NOAEL threshold (No Observed Adverse Effect Level)- the largest amount of a substance that the most sensitive animal model can consume without noticeable adverse health effects. This threshold approach is more and more criticized because it doesn't guarantee the lack of risks to every individual from a population, and new science-based mechanism models are developed, that give insight to physiology and pharmacokinetics. Moreover,

thresholds are not suitable for genotoxic carcinogens because a single molecule can initiate a cancer – Tjihuis (2012). In epidemiology, hidden associations between real human exposures and health effects (benefits or adverse effects) are researched and quantified.

On the other hand, benefit assessment is a new field, that mimics risk assessment, within the field of nutrition. Nutritional challenge tests are developed to measure a system’s (i.e. human) physiologic robustness and to discriminate between individuals, and therefore to orient benefit-risk assessment towards a more individ-oriented approach -Tjihuis (2012) (see also Section 3). It is considered, in general, that benefit assessment should mirror the traditional steps of risk assessment, with a benefit arm and a risk arm running in parallel until final comparison (Figure 1.1). Some projects (BRAFO, QALIBRA, PASSCLAIM) follow this approach, starting with a qualitative tier that tries to inference a net benefit or a net disadvantage, and ending up with a quantitative measure (like QALY -Quality-Adjusted Life Years- see Section 2), if that was not possible.

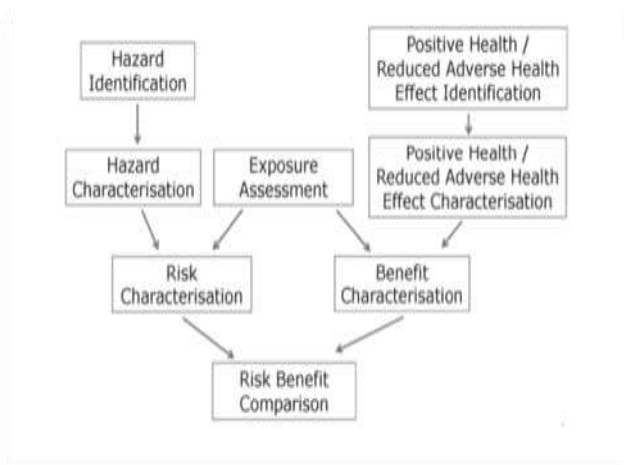


Fig. 1.1. Guidance on human health risk-benefit assessment of foods -EFSA (2010)

The BENERIS project takes a step further to acknowledge the complexity, uncertainty and openness of the task of risk-benefit assessment by introducing the notion of “open assessment”: the objective of each tier in the system is to convince a

critical outsider (open criticism) about conclusions. Outsiders can join and criticize the current content. The main tools used in BENERIS are probability distributions and Bayesian nets (<http://en.opasnet.org/w/Beneris>). By their intrinsic nature, argumentation systems are in our opinion by far more suitable to model such non-monotonic evolving informational systems, with knowledge that can be defeated or re-instated by new pieces of information (Sections 4, 5).

2.QALYs, DALYs AND THEIR LIMITATIONS

QALYs (Quality-Adjusted Life Years) are measures defined in the 1970s to improve decision making on expenditures in healthcare, and are currently the most used integrated health measure in food-related benefit-risk analysis.

QALY (Figure 2.1) is a composite metric that takes into account quality of life (morbidity) and survival (mortality):

$$(2.1) \text{ QALY} = \text{Time (duration of disease in years)} \times \text{Utility (quality of life } q) + \text{age of onset of disease (AoO: a number which is lower than the true age of death- AoD)}$$

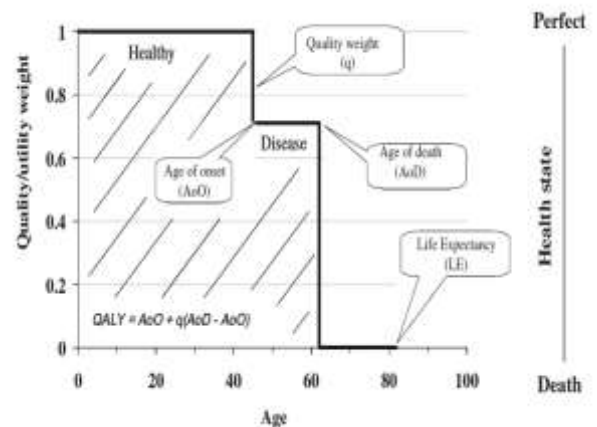


Fig. 2.1. Schematic representation of the QALY concept (one individual accounting for one episode of disease) -from Tjihuis (2012).

As we can see in Figure 2.1 full health is represented by weight 1, death has weight 0, and each disease decreases the state of full health to a weight “q” between 0 and 1:

$$(2.2) q=1-w,$$

and w is a disability weight that mirrors the subjective perception of different populations over disorders (taken usually from the WHO website: Global Burden of Disease 2010 study, e.g. for Angina pectoris w=0.124)

QALYs are suitable to measure health gains at micro-scale (to compare 2 interventions) for an individual or to compare the same intervention for two individuals (Figure 2.2).

But quantifying the quality of life is a challenging and controversial idea: different populations evaluate conditions differently, and valuing one individual’s life over another’s is also a source of mistakes (being unethical). For instance, who could decide that a shorter full-health life is to be preferred over a longer wheel-chaired life?

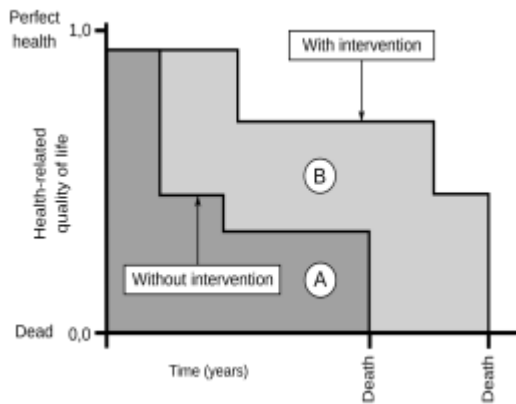


Fig. 2.2. Demonstration of quality-adjusted life years (QALYs) for two individuals. Individual A (who did not receive an intervention) has fewer QALYs than individual B (who received an intervention). (from Qalibra project site: www.qalibra.eu)

The main drawback of a QALY is its link to healthcare funding, raising serious ethical concerns.

“...the quantitative nature of QALYs reflects an overly utilitarian approach.” -Pettitt (2016). Moreover, “a generic list of QALYs reduces the role and the expertise of healthcare providers and ultimately undermines their ability to make judgments based on an individual’s need” -Pettitt (2016). That is because in order to take a decision (when comparing 2 treatment groups, for instance), QALYs are summed over all individuals of a population and the intervention with the greater average areas under the curve is chosen (considered to give the highest health maximization per population) - as such, being not individ -oriented).

$$(2.3) DALY= \text{Years lost (YL)} + \text{years lived with the disease} * \text{disease burden } w$$

Therefore, the lowest DALY gives the highest health maximization. DALYs discriminate the elderly and the disabled populations, those with mental illnesses and cancer patients -Tijhuis (2012). For instance, old or cancerous people are disadvantaged by this calculus because having fewer (statistical supposition) years left to live, their contribution to the sum is smaller.

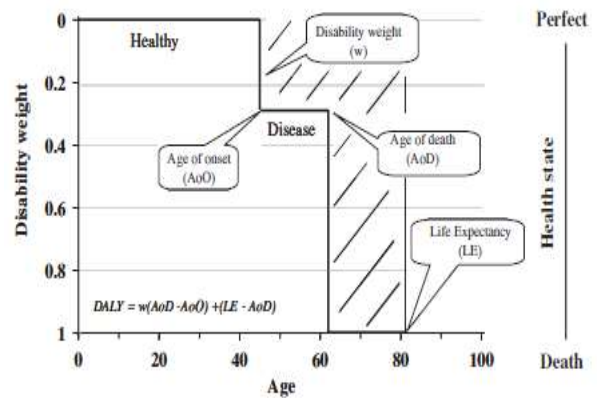


Fig. 2.3. Calculation using disability-adjusted life years (DALYS) is equivalent to QALY, but with the vertical scale reversed, so that 0 represents full health and 1 represents death. (www.qalibra.eu)

As we can see, the choice for DALY/QALY is rather a pragmatic one (based on experience), not a fundamental one.

As conventional medicine is currently evolving from the “one-size-fits-all” approach towards personalized medicine, that makes prognosis and adjusts therapies according to quickly developing patient-oriented metrics that stem from genetics and genomics, the idea of adjusting QALYs to personalized parameters becomes more obvious. One such attempt is the “window of benefit” -Palou (2009).

The window of benefit is a framework that combines thresholds and scores in order to define an intermediate interval of optimal individual benefit: from LLAB (lower level of additional benefit) to ULAB (upper level of additional benefit), ranging in the simplest cases between the known limits of RDA (recommended dietary allowance-intake sufficient to prevent deficiency) and UL (upper intake level= avoid toxicity). (Figure 2.4)

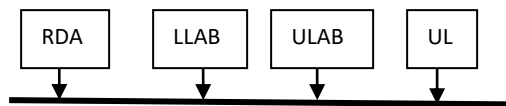


Fig. 2.4. From RDA to UL

In general, the threshold approach computes a limit that must not be exceeded for toxins or that must be reached for nutrients.

RDA and UL are derived from the NOAELs in rodents, assuring maximal safety under the toxicological view, but they do not assure optimal health form the nutritional point of view. Moreover, risk assessment at present determines the percentage of population that exceeds UL (plus magnitude and duration of excessive intake), but does not compute the individual risk of each person.

To personalize the decision, upper and lower bounds of benefits must be established (ideally considering age, gender, lifestyle and genotype), similarly to the thresholds for risks, such that an intake between LLAB and ULAB is considered protective against a specified health/nutritional benefit.

The most important aspect is that this individual window (LLAB-ULAB) shifts depending on age, sex, genetic constitution and lifestyle. (Figure 2.5)

An important observation is that if a certain nutrient is related to multiple risks/ benefits, an LLAB-ULAB window must be defined separately for each of them, and it may happen that the minimal dose for a certain benefit (e.g. reducing cancer) is higher than the minimal dose for an adverse effect (e.g. estrogenic hormone disruption).

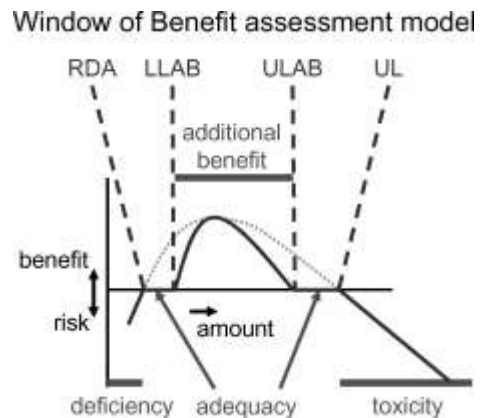


Fig. 2.5 “Schematic representation of the Window of Benefit assessment model. An increase in the intake of a food or food component will result in the following subsequent physiological conditions: deficient, adequate, beneficial, adequate, and toxic....” –from Palou (2009)

Sometimes, a small adverse effect could be accepted because of its larger beneficial effect at a higher dose. As such, LLAB and ULAB will not always be within the range defined by RDA or UL. Also, an ordering of adverse effects, according to their severity and type is needed, dependent on individual parameters (gender, age, genotype). For the window of benefit approach, the authors suggest combining the various limits for benefits and adverse effects, classified after their type and severity, and computing the LLAB-ULAB of optimal intake through weighted decision. We will consider a different model in Section 5, using Valued Arguments for non-monotonic reasoning.

Ordering benefits according to their value is strongly related to the score model used by the window of benefit. The score model implies defining a generic tool to score health-related benefits, that calculates notes for every nutrient and combines these notes into a global score.

To build a scientifically valid evaluation, a mechanism-based approach systems biology understanding of health effects (stemming from cell physiology, molecular biology and biochemistry) is needed. This approach is also useful to see if unintended effects occur, to weighing whole-foods, that have a combined effect on our health (because they have different components), and to define early biomarkers of effect (sensitive markers to identify deviation from homeostasis, that are indicative of a later pathological effect). These biomarkers can be used to assess if a food product is good for everyone or only for a certain age, gender and genotype defined subgroup. It is known that genetic polymorphisms can produce a phenotype more sensitive to specific adverse effects/ health matters.

For instance, MTHFR (Methylene-Tetra-Hydro-Folate-Reductase) is a gene that controls the homocysteine levels -Trimmer (2013), which, in turn, are very important for the production of glutathione- the body's most important antioxidant which plays a major role within detoxification of harmful, disease-causing toxins. As researched by Pu (2013), a link exists between the mutations of this gene and the risk of developing autism in children exposed to different toxicities. We shall use this category of mutations for the Example in Section 6.

3. NUTRIGENOMICS AND HEALTH BIOMARKERS

Nutrigenomics is an evolving field that aims to develop personalized nutrition. Genes regulate our metabolism, our susceptibility to certain diseases, our response to nutrients and our resistance to toxic substances. As such, individuals have different windows of benefit in terms of nutrient intake, as we have seen above. As stated by Hesketh (2006), “specific nutrient intake based on the whole population may have a risk for some individuals, a benefit for others and have little effect in others”.

The sequencing of the human genome plus the single nucleotide polymorphisms (SNPs) that express individual genetic variations and the rapid technologies to analyze them make it possible in the present to answer the question “to what extent can we identify sub-groups within populations, and which

genetic factors provide them with increased benefit or increased risk from a particular nutrient at a certain intake?” –Hesketh (2006)

If we consider nutrition- single nucleotide polymorphism interactions, it is already known that a certain benefit/ risk are the effect of a specific interaction between a nutrient and the genotype (for instance, increased PUFA intake reduces HDL-C only among homozygous -GG- for apoA1-75 (G/A) SNP i.e. 75% of the population, and it increases HDL-C in all the others). But to make things even more complex, different SNPs interact, so we need information on SNPs in multiple genes to decide the benefit or risk of an individual. Moreover, lifestyle and environmental differences must be accounted for beyond the genetic influences. By now, no clear approach exists to deal with this -Hesketh (2006), and we think the complexity of the problem makes it a perfect candidate for practical reasoning (uncertainty, incomplete information, non-monotonicity: interaction between multiple SNPs can contradict conclusions for sub-groups) (Section 5).

Nutritional-relevant SNPs are still to be discovered. More limitations arise from the fact that early biomarkers of risk are difficult to define, as such it is still questionable how to define optimal health (a “healthy phenotype”).

Research in the field - Elliot (2007) suggests that health is not, simply, the absence of any known disease, but should rather be defined as the “ability to withstand and manage relevant physiological challenges so that homeostasis is maintained” (some relevant tests already exist that support this, like the oral glucose test to assess glycaemic control or the effort stress test for cardiac conditions). Current challenge tests should be enriched by the functional genomic technologies (transcriptomics, proteomics, metabolomics) to develop new biomarkers of health. Optimal biomarkers should not only be robust, practical and sensitive, but should also be mechanism-linked, so that implications of changes can be understood, because “unexpected biological effects and interactions plus inter-individual differences in

nutritional requirements” hinder rigorous benefit-risk analysis.

Elliot & all. (2007) argue that the “omic” approaches can help to describe entire biological responses (i.e. patterns of parameters from integrated multi-component biomarkers) and the rate and completeness of the return to the original status, while being also useful to define dynamic measures (i.e. not measurements at a single point, but fluctuations in time). The ultimate goal of these biomarkers is to determine the “earliest possible indications of long-term disturbances” –Elliot (2007). Research is still to be done to identify these diet-related relevant challenge tests and their integration with the “omics” technologies, such that the result will be suitable for the complexity of foods.

4. VALUED ARGUMENTATION FRAMEWORKS

We believe more attention should be paid to the qualitative comparison of risks and benefits, and we argue that valued argumentation frameworks are a suitable approach that encompasses the non-monotonicity, complexity, incompleteness, uncertainty and individ-dependency of the task.

Valued Argumentation Frameworks are standard Argumentation Frameworks –see Dung (1995), that add values to arguments by means of a function:

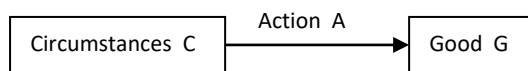
Definition 1. (Trevor (2003)). A value-based argumentation framework (VAF) is a 5-tuple:

$$(4.1) \text{VAF} = \langle \text{AR}, \text{attacks}, V, \text{val}, P \rangle,$$

where AR is a finite set of arguments, *attacks* is a non-reflexive, binary relation on AR, V is a nonempty set of values, *val* a function which maps from AR to V and P are the possible audiences individuated by their different preferences over values.

While standard argumentation frameworks are suitable for reasoning about matter of facts, their valued enhancement was necessary for practical reasoning, where arguments have often the form - Trevor (2003) :

Arg 1:



which reads “Action A should be performed in circumstances C, because the performance of A in C would promote some good G.”-Trevor (2003)

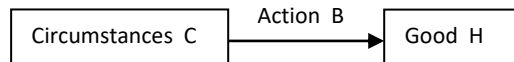
This kind of arguments can be attacked in 3 (known) ways:

1. not obtaining C (false premises)
2. Action A in C does not promote G (conclusion doesn't follow the premises)
3. Action B in C promotes G, and B excludes A,

But two new types of attacks can be considered:

4. G is not “good” (re-evaluation of goods)
5. there is a *more desirable* good H that can be obtained by action B that excludes A:

Arg 2:



So, the ends are ordered according to a preference relation, and different circumstances may lead to different preferences over the ends.

Definition 2. (Trevor (2003)). An *audience-specific value-based argumentation framework* (AVAF) is a 5-tuple:

$$(4.2) \text{VAF}_a = \langle \text{AR}, \text{attacks}, V, \text{val}, \text{valpref}_a \rangle$$

where *a* is an audience and $\text{valpref}_a \subseteq V \times V$ is a preference relation (transitive, non-reflexive, asymmetric), and v_1 is preferred to v_2 is written as $\text{valpref}_a(v_1, v_2)$.

Therefore, one and the same argument can have different values in different circumstances (that generate different preference relations).

Definition 3. (Trevor (2003)). An argument $A \in \text{AF}$ *defeats_a* an argument $B \in \text{AF}$ for audience *a* iff both $\text{attacks}(A, B)$ and not $\text{valpref}_a(\text{val}(B), \text{val}(A))$.

Definition 4. (Trevor (2003)). An argument $A \in AR$ is *acceptable to audience a* ($acceptable_a$) with respect to set of arguments S , ($acceptable_a(A,S)$) if:

$$(4.3) (\forall x)((x \in AR \& defeats_a(x,A)) \rightarrow (\exists y)((y \in S) \& defeats_a(y,x)))$$

Therefore S represents the defenses of an attacked argument.

Definition 5. (Trevor (2003)). A set of arguments is *conflict-free for audience a* if

$$(4.4) (\forall x)((x \in S \& y \in S) \rightarrow (\neg attacks(x,y) \vee valpref_a(val(y),val(x)) \in valpref_a))$$

The notion of conflict-free sets of arguments stands for coherence (not supporting conflicting facts).

Definition 6. (Trevor (2003)). A conflict-free for audience a set of arguments S is *admissible for an audience a* if

$$(4.5) (\forall x)(x \in S \rightarrow acceptable_a(x,S).$$

Definition 7. (Trevor (2003)). A set of arguments S in a value-based argumentation framework is a *preferred extension for audience a* ($preferred_a$) if it is a maximal (with respect to set inclusion) admissible set of AR for audience a (i.e. a maximal coherent point of view).

Note that for a given preference over values $valpref_a$ we can break the cycles that contain different values by removing those attacks faced with superior values ($val(A_j) > val(A_i)$)- Figure 4.1.

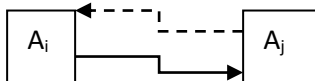


Fig. 4.1. Breaking attack cycles ($val(A_j) > val(A_i)$)

4.1. Acceptance in value-based argument frameworks

Theorem. (Trevor (2003)) Every AVAF with no single-valued cycles has a unique, nonempty preferred extension (i.e. a unique non-empty coherent conclusion of argumentation).

An algorithm linear in the number of attacks in the AF exists for AVAF's with no single-valued cycles (EXTEND -Trevor (2003)). The author argues, moreover, that single-valued cycles are a sign of flawed reasoning (odd cycle: nothing can be believed; even-cycles: a dilemma that requires a choice between alternatives), and should be resolved before a practical argumentation session.

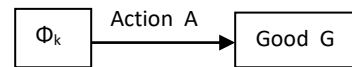
5. VALUED ARGUMENTS IN INDIVIDUALISED BENEFIT-RISK ASSESSMENTS OF FOODS

Suppose N SNP-related relevant phenotypes have been defined:

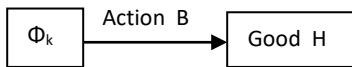
$$(5.1) \Phi = \{\Phi_1, \dots, \Phi_N\} \text{ (these describe } N \text{ known population sub-groups)}$$

Also consider the following arguments (model 4.1):

Arg 1:



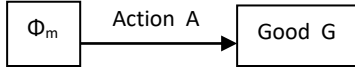
Arg 2:



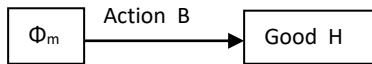
that are in an attack relation - the 5th type of attack (there is a *more desirable* good H that can be obtained by action B that excludes A), and Φ_k describes the circumstances of the k^{th} phenotype. These particular circumstances also determine a certain ordering of goods, such that H is more desirable than G for phenotype Φ_k , but not necessarily in general. This means that *Arg 2* defeats *Arg 1*, if *Arg 2* is not further attacked by other *Facts* or *Arguments* relevant for an individual belonging to this phenotype. For

another phenotype Φ_m , we might have the same actions and goods but a different ordering:

Arg 3:



Arg 4:



such that G is preferable to H in this context- Φ_m . So Arg 4 defeats Arg 3, if no further evidence or argument reinstates it.

Therefore, we obtain a Valued Argumentation Framework:

$$(5.2) \langle AR, attacks, V, val, \Phi \rangle,$$

where Φ is the set of phenotypes, that define the relative importance of different goods (i.e. nutritional benefits for health).

Practical reasoning within this framework is performed with algorithm EXTEND (Section 4).

6.EXAMPLE

Ginsberg (2009) weighs the benefits and the risks of fish consumption across different species. Shortly, while omega 3 fatty acids from fish are very good for both neurodevelopment and cardiovascular health, methyl-mercury, which is a non-avoidable contaminant, impairs both these health aspects. As the concentration of the nutrient and of the contaminant vary along different species, the paper analyses which are the best-compromise fish options.

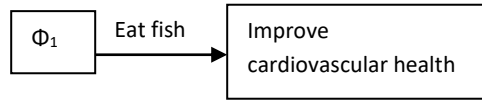
While many phenotypes detoxify small methyl-mercury amounts well, the carriers of the MTHFR mutation have a problem, as we have discussed in the end of Section 2.

We shall consider two population subgroups: Φ_1 - the carriers of MTFHR mutations and Φ_2 - the non-

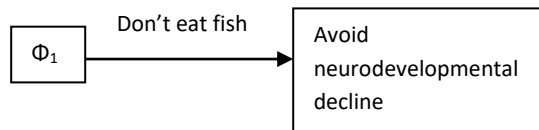
carriers, and we shall model benefit –risk assessment reasoning about eating fish, using the model defined in Section 5.

Φ_1 : MTFHR mutations

A 1':

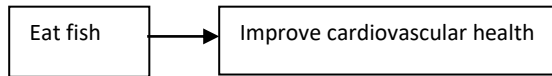


A 2':

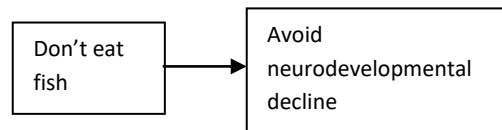


We shall simplify A 1' and A 2' to A 1 and A 2, respectively:

A 1:



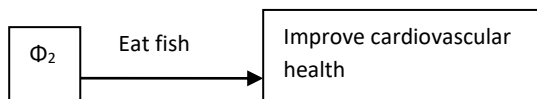
A 2:



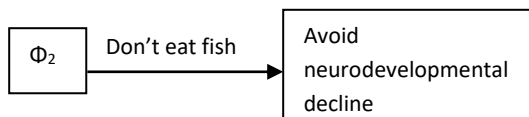
and Φ_1 is the context which assigns a specific value to them: $val(A1)=v1$, $val(A2)=v2$ and $valpref_{\Phi_1}(v2, v1)$. Therefore, as A 1 and A 2 attack each other (mutually exclusive actions), A 2 defeats A 1 within context Φ_1 .

Φ_2 : no MTFHR mutations

A 3:

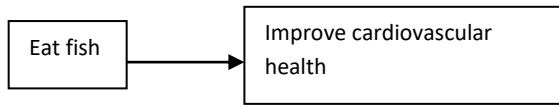


A 4:

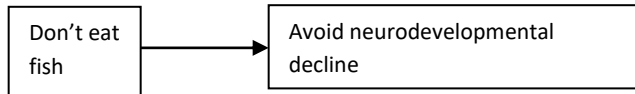


which also can be simplified to:

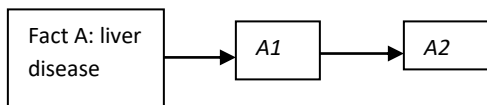
A 1:



A 2:



Here, $\text{val}(A1) = v1$, $\text{val}(A2) = v2$, but $A1$ defeats $A2$ because $\text{valpref}_{\phi_2}(v1, v2)$. This is valid, of course if $A1$ is not itself attacked by other pieces of evidence, like facts relevant to an individual. For instance, if a non-carrier of MTFHR mutations happens to have a liver disorder which affects its detoxification function, we have the following reasoning chain:



If $\text{val}(\text{Fact A}) = v3$, and $\text{valpref}_{\phi_2}(v3, v1)$, then $A2$ is reinstated and, as such, eating fish is not a good option, even if we refer to a non-carrier of MTFHR mutations.

7. CONCLUSIONS AND FUTURE WORK

The paper introduces the challenges of benefit risk-assessment of foods and drugs, faced with modern trends of personalized medicine. We argue that a practical reasoning framework suits the problem better than other approaches and present an example to illustrate the idea. As new biomarkers of health will be developed, and more insight will be gain about nutrient/ contaminant bio-mechanics within different phenotypes, the theoretical background should be adapted and adjusted to tailor the problem.

7. REFERENCES

- P. M. Dung (1995). "On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n-person games". *Artificial Intelligence*. **77** (2): 321–357.
- EFSA (2010) *EFSA Journal* 2010; 8(7):1673;
- R. Elliot, C. Pico & all. (2007) "Nutrigenomic approaches for benefit-risk assessment analysis of foods and food components: defining markers of health", *British Journal of Nutrition* (2007), 98, 1095-1100;
- G. Ginsberg, B. Toal, (2009) "Quantitative Approach for Incorporating Methylmercury Risks and Omega-3 Fatty Acid Benefits in Developing Species-Specific Fish Consumption Advice", *Environmental Health Perspectives*, 2009, 117 (2): 267-275;
- J. Hesketh, I. Wybranska et all. (2006) "Nutrient-gene interactions in benefit-risk analysis", *British Journal of Nutrition* (2006), 95, 1232-1236;
- A. Palou, C. Pico, (2009) "Integration of Risk and Benefit Analysis—The Window of Benefit as a New Tool?", *Critical reviews in food science and nutrition*, September 2009
- D.Pettitt & all (2016) "The Limitations of QALY: A Literature Review", *Journal of Stem Cell Research*, 2016
- Pu D, Shen Y, Wu J. (2013) Association between MTHFR gene polymorphisms and the risk of autism spectrum disorders: a meta-analysis. *Autism Res*. 2013 Oct;6(5):384-92. doi: 10.1002/aur.1300.
- M.J. Tjshuis, & all. (2012) "State of the art in benefit-risk analysis: Food and nutrition" , *Food and Chemical Toxicology* 50 (2012), 5-25;
- Trevor J.M, Bench-Capon (2003) "Persuasion in Practical argument Using Value-based Argumentation frameworks", *J. of Computational Logic*, V 13, no 3, Oxford University Press, 2003
- Trimmer EE. (2013) Methylene tetrahydrofolate reductase: biochemical characterization and medical significance. *Current Pharmaceutical Des*. 2013;19(14):2574-93.