The Identity and Mereology of Pathological Dispositions

Adrien Barton[1,*], Olivier Grenier[1] & Jean-François Ethier[1,*]
[1] GRIIS, Université de Sherbrooke
Sherbrooke, Quebec, Canada
*adrien.barton@gmail.com; ethierj@gmail.com

***Abstract*—Diseases, risks of pathological processes and predispositions have been formalized as dispositions. The relations between those pathological dispositions, however, remain unclear. We apply here a recently developed theory of mereology and identity among dispositions to analyze such relations. In particular, we show how a framework for the identity of dispositions leads to a disease being realized not only by its disease course but also by each of its pathological process; how it avoids risk multiplicativism; and how a predisposition can be identified with a risk whose estimated probability is higher than the probability of the risk for a reference class. We discuss how this makes a predisposition always relative to a reference class, to a time-frame and to sources of risk estimates; and we clarify the nature of risk factors.**

Keywords—disposition; disease; risk; predisposition; risk factor; identity; mereology

I. Introduction

Dispositions are entities such as fragility, solubility, or vulnerability to poison, which can be triggered by some process, leading to a realization process. For example, the fragility of a glass might be triggered by a shock and realized by the process of this glass breaking. Dispositions may exist even if they are not realized or even triggered: this glass if fragile even if it never breaks, or even if it never undergoes any shock. Dispositions are omnipresent in the clinical domain. An obvious example concerns predispositions to diseases. But other entities have also been formalized as dispositions, such as diseases themselves (by the Ontology for General Medical Science, OGMS [1]), and medical risks [2]. The relations between various dispositions is not always easy to determine. For example: is the disposition realized by a disease course identical to the disposition realized by each of the pathological processes that are parts of this disease course? What is the connection between the risk to develop a disease and this disease? Is the risk to get a stroke over 6 months the same entity as the risk to get a stroke over 12 months? What is the connection between a predisposition to a disease and the risk to develop this disease? And the connection between such a predisposition and risk factors for this disease? Recently, ontological frameworks have been proposed for the identity [3] and mereology [4] of dispositions. This article will show how those frameworks can clarify the nature of such dispositions and thus answer the above-mentioned questions. The formalization will be written using the Manchester Syntax [5] for OWL [6] with occasional use of first-order logic. Particulars and relations will be written in bold, and classes in italic.

II. The Identity and Mereology of Dispositions

Reference [3] presents two frameworks for the identity of dispositions. In the first framework, called ONLY, a particular disposition **d** is said to have a realization specification *R* and a trigger specification *TR* (abbreviated in the remainder of this paper "disposition to *R* when *TR*") if the following relations hold:

d has_trigger only *TR*

d has_realization only *R*

That is, all triggers (resp. realizations) are instances of a same specific class. On the other hand, in the PARTHOOD framework, any process that has as part a trigger of **d** is also a trigger of **d**, and any process that is a part of a realization of **d** is also a realization of **d**. This leads to define the class of "minimal triggers" of **d**: the class of triggers of **d** of which no proper part is a trigger of **d**; and the class of "maximal realizations" of **d**: the class of realizations of **d** which are not proper parts of another realization of **d**. For example, a strong shock on a glass would be a trigger of the glass' fragility (which is a disposition), whereas a minimal trigger of the fragility would be the part of this strong shock during which it undergoes a critical pressure that makes it break; and the maximal realization of this fragility would be the glass breaking, while a part of this process would also be a realization, though not maximal.

This paper also introduced, for all practical purposes, a criterion of identity (named here "ID") acceptable for all practical purposes in both ONLY and PARTHOOD, which states that two dispositions **d** and **d'** are identical iff they have the same categorical basis, the same class of triggers and the same class of realizations; using relations as defined in [7] , we can write:

(ID) **d** is identical to **d'** iff [(∃**cat**, **d has_basis cat** ∧ **d' has_basis cat**) ∧ (**trigger_of** value **d**) EquivalentTo (**trigger_of** value **d'**) ∧ (**realization_of** value **d**) EquivalentTo (**realization_of** value **d'**)]

Reference [4] presents a theory of mereology among dispositions. It distinguishes in particular two kinds of mereological relations among dispositions named "**mod-part_of**" and "**add-part_of**". For example, the disposition to attract another magnet when facing an unlike pole and the disposition to repulse the very same magnet when facing a like pole are mod-parts of a magnet's ferromagnetic disposition. And the disposition to dissolve of the left half of a tablet and the disposition to dissolve of the right half of this tablet are add-parts of the whole tablet's disposition to dissolve.

Reference [4] also presents several axioms satisfied by those dispositions. The bearer of a disposition-part (whether mod-part or add-part) is always a part of the bearer of the disposition-whole. A mod-complex (that is, a disposition that has a proper mod-part) is triggered by a process if and only if at least one of its proper mod-parts is triggered by this process; and it is realized in a process if and only if at least one of its proper mod-parts is realized in this process. If an add-complex (that is, a disposition that has a proper add-part) is triggered by a process, then all its add-parts are triggered by a part of this process; and if it is realized in a process, then all its add-parts are realized in a part of this process.

We will apply those theories of identity and mereology to several questions: is a disease realized by its disease course, by each of its pathological processes, or by all of those? Are several risks the same entity or different ones? And finally, what are the connections between a predisposition to a disease, the general risk to this disease, and the risk factors of this disease?

III. Diseases

Consider Mr. Miller, who is epileptic. Diseases are formalized in OGMS as dispositions realized by various pathological processes that are parts of a disease course. Suppose that over the course of his life, Miller undergoes n epileptic crisis: **crisis$_1$**, …, **crisis$_n$**. The epileptic disease course of Miller has as parts all those epileptic crises (and possibly pathological processes of other types, though we will ignore them here): **epileptic_dc$_0$ has_part crisis$_i$** for every i between 1 and n. Suppose moreover that Miller is a photosensitive epileptic, and that for every i between 1 and n, a flash of light **flash$_i$** triggers his **crisis$_i$**. Consider now how each of the two ontological frameworks would formalize this situation.

A. Analysis in ONLY

In ONLY, one could distinguish two dispositions borne by Miller, that might each be identified with the epileptic disease: **epilepsy$_0$**, which is realized once (over an extended time period) by Miller's whole disease course, and **epilepsy$_1$**, which is realized multiple times by each of Miller's individual epileptic crisis:

epilepsy$_0$ has_realization only *Epileptic_disease_course*

epilepsy$_1$ has_realization only *Epileptic_crisis*

More specifically:

epilepsy$_0$ has_realization epileptic_dc$_0$

epilepsy$_1$ has_realization crisis$_i$ (for i between 1 and n)

epilepsy$_0$ and **epilepsy$_1$** are different, as they do not have the same realizations. More specifically, **epilepsy$_1$** is a proper add-part of **epilepsy$_0$**. Indeed:

- the bearer of **epilepsy$_0$** is the same as the bearer of **epilepsy$_1$**.
- any trigger of **epilepsy$_1$** (a light flash) is a part of the trigger of **epilepsy$_0$** (the mereological sum of all light flashes).
- any realization of **epilepsy$_1$** (a pathological process of epileptic crisis) is a part of the realization of **epilepsy$_0$** (the disease course of the epilepsy).

Those relations are represented below on Figure 1.

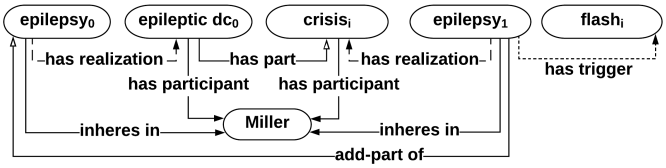

Fig. 1. A formalization of disease in ONLY

The problem of this formalization is that it artifically introduces two dispositions **epilepsy$_0$** and **epilepsy$_1$**. These two dispositions have the same categorical basis: the qualities that are responsible for the realization of the epileptic crises are exactly those that are responsible for the realization of the mereological sum of those epileptic crises. However, the medical science does not seem to distinguish those two different dispositions, but would rather identify them. A more economical option would therefore be to introduce an alternative disposition **epilepsy$_2$** which is realized by the disease course, but also by each pathological process. We will now show that this is enabled by the second ontological framewok PARTHOOD.

B. Analysis in PARTHOOD

In PARTHOOD, any part of a realization of a disposition **d** is also a realization of this disposition. Therefore, if we postulate that **epilepsy$_2$ has_realization epileptic_dc$_0$**, then for every i between 1 and n, **crisis$_i$** is also a realization of **epilepsy$_2$**: **epilepsy$_2$ has_realization crisis$_i$**.

Thus, we can state, for every i between 1 and n:

- **epilepsy$_2$ has_trigger flash$_i$**
- **epilepsy$_2$ has_realization crisis$_i$**
- **epilepsy$_2$ has_realization epileptic_dc$_0$**

We can also add that Miller's epileptic disease course is a maximal realization of **epilepsy$_2$**. These relations involving particulars can be generalized into relations between classes as represented in the figure 2 below (using relations defined in [3] and [7]).

Note that in PARTHOOD, one cannot say that '*Epilepsy* **has_trigger** only *Flash*' – because any process that has a *Flash* as a part is also a trigger of the epileptic disease; however, we can state an axiom concerning a subclass of *Flash* named *Epilepsy-triggering Flash*, namely:

Epilepsy-triggering Flash **trigger_of** some *Epilepsy*

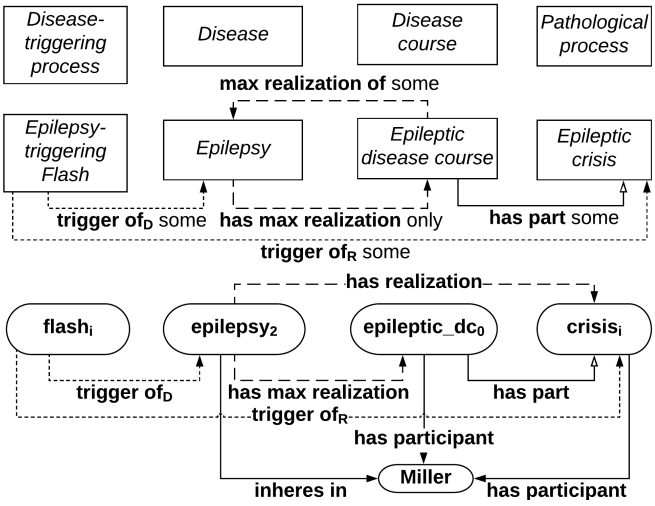

Fig. 2. A formalization of disease in PARTHOOD

Such an ontological framework seems preferable: we do not have to duplicate artificially the disease into a disposition realized by the disease course and a disposition realized by each pathological process.

IV. RISKS

In [2], a risk of stroke for people with atrial fibrillation was formalized as a disposition $Risk_{AF,Stroke}$; the risk of stroke for people with atrial fibrillation over 12 months was formalized as another disposition $Risk_{AF,12m,Stroke}$; and the risk of stroke over 6 months was formalized as yet another disposition $Risk_{AF,6m,Stroke}$. The connection between $Risk_{AF,Stroke}$, $Risk_{AF,12m,Stroke}$ and $Risk_{AF,6m,Stroke}$ was left open. We will analyze it in the two frameworks of identity of dispositions.

A. Analysis in ONLY

The framework ONLY raises two issues for the formalization of risks developed in [2]. The first one is that the trigger specification of $Risk_{AF,12m,Stroke}$ is the class of 12-months-long history-parts (abreviated $H\text{-}P_{12m}$) of its bearer. Thus:

$Risk_{AF,12m,Stroke}$ **has_trigger** only $H\text{-}P_{12m}$

However, it is not the case that everything that happens in the body of a person has a causal influence on the possible realization of his risk. Therefore, some proper parts of instances of $H\text{-}P_{12m}$ (the ones that are causally relevant) should also be triggers of $Risk_{AF,12m,Stroke}$.

The second problem is that $Risk_{AF,Stroke}$, $Risk_{AF,6m,Stroke}$ and $Risk_{AF,12m,Stroke}$ cannot be identical with each others. $Risk_{AF,Stroke}$ is triggered by all history-parts of its bearer (that is, all parts of his history – which is the mereological sum of all processes taking place in or on his body [8]). $Risk_{AF,12m,Stroke}$ is only triggered by 12-months-long history-parts (abreviated $H\text{-}P_{12m}$) of its bearer, and $Risk_{AF,6m,Stroke}$ by 6-months-long history-parts.

Consider a patient Jones who has AF, and his risks $\mathbf{risk}_{\mathbf{Jones},Stroke}$, $\mathbf{risk}_{\mathbf{Jones},12m,Stroke}$, and $\mathbf{risk}_{\mathbf{Jones},6m,Stroke}$ (as defined by [2]). Then, $\mathbf{risk}_{\mathbf{Jones},6m,Stroke}$ is an add-part of $\mathbf{risk}_{\mathbf{Jones},12m,Stroke}$; as a matter of fact:

- the bearer of $\mathbf{risk}_{\mathbf{Jones},Stroke}$, $\mathbf{risk}_{\mathbf{Jones},12m,Stroke}$, and $\mathbf{risk}_{\mathbf{Jones},6m,Stroke}$ is the same, namely **Jones**.
- any trigger of $\mathbf{risk}_{\mathbf{Jones},12m,Stroke}$ is a 12 months-long history part of Jones, and thus has as part at least one (and actually many, if not an infinity) 6 months-long history part of Jones, which is a trigger of $\mathbf{risk}_{\mathbf{Jones},6m,Stroke}$
- any realization of $\mathbf{risk}_{\mathbf{Jones},12m,Stroke}$ is a realization of $\mathbf{risk}_{\mathbf{Jones},6m,Stroke}$ (e.g. the realization on December 31st 2018 of Jones' 12-months risk of stroke between January 1st and December 31st 2018 is also the realization of Jones' 6-months risk of stroke between July 1st 2018 and December 31st 2018)

Moreover, both $\mathbf{risk}_{\mathbf{Jones},6m,Stroke}$ and $\mathbf{risk}_{\mathbf{Jones},12m,Stroke}$ are mod-parts of $\mathbf{risk}_{\mathbf{Jones},Stroke}$ (that is, Jones' 6-months risk of stroke and 12-months risk of stroke are two of the many modes of realization of his general risk of stroke); and indeed:

- the bearer of $\mathbf{risk}_{\mathbf{Jones},Stroke}$, $\mathbf{risk}_{\mathbf{Jones},12m,Stroke}$, and $\mathbf{risk}_{\mathbf{Jones},6m,Stroke}$ is the same, namely **Jones**.
- any trigger of $\mathbf{risk}_{\mathbf{Jones},12m,Stroke}$ or $\mathbf{risk}_{\mathbf{Jones},6m,Stroke}$ is a trigger of $\mathbf{risk}_{\mathbf{Jones},Stroke}$.
- any realization of $\mathbf{risk}_{\mathbf{Jones},12m,Stroke}$ or $\mathbf{risk}_{\mathbf{Jones},6m,Stroke}$ is a realization of $\mathbf{risk}_{\mathbf{Jones},Stroke}$.

Finally, one can introduce various risk estimates, such as $\mathbf{risk\_estimate}_{\mathbf{Jones},12m,Stroke}$, which **is_about** $\mathbf{risk}_{\mathbf{Jones},12m,Stroke}$ (alternatively, we can introduce the relation **object_of** inverse of **is_about**, and state that $\mathbf{risk}_{\mathbf{Jones},12m,Stroke}$ is the **object_of** $\mathbf{risk\_estimate}_{\mathbf{Jones},12m,Stroke}$).

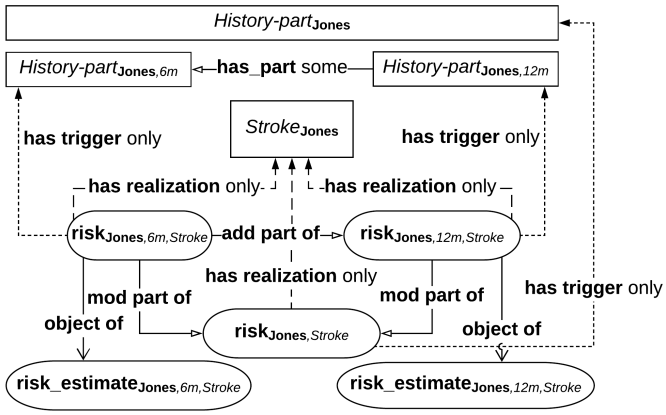

Fig. 3. A formalization of risk of stroke and estimates in ONLY

Thus, the relations between the various risks can be described by the theory of mereology of dispositions. However, this leads to risk multiplicativism: for any class of time intervals with a given length of time t, there is a corresponding disposition $\mathbf{risk}_{\mathbf{Jones},t,Stroke}$.

B. Analysis in PARTHOOD

The framework PARTHOOD helps to solve the two issues mentioned above. As a matter of fact, in PARTHOOD, a disposition does not necessarily have only one class of trigger (or one class of realization). Thus, it does not exclude that proper parts of instances of $H\text{-}P_{12m}$ (in particular, the causally

relevant parts) would be triggers of $Risk_{AF,12m,Stroke}$. Moreover, in PARTHOOD, any process that has as part a trigger is also a trigger. Thus, if a particular 6-months-long history part **h-p6** is a trigger of a disposition, then a particular 12-months-long history part **h-p12** that has **h-p6** as part is also a trigger of this disposition. As a consequence, we cannot define as earlier entities such as $\mathbf{risk}_{\mathbf{Jones},6m,Stroke}$ or $\mathbf{risk}_{\mathbf{Jones},12m,Stroke}$. However, we can still define $\mathbf{risk}_{\mathbf{Jones},Stroke}$, which can be triggered (in particular, but not only) by history-parts of Jones of any time length. To relate $\mathbf{risk\_estimate}_{\mathbf{Jones},6m,Stroke}$ with its 6-months time frame, we can assert it as being not only about $\mathbf{risk}_{\mathbf{Jones},Stroke}$, but also about the set of triggers $H\text{-}P_{\mathbf{Jones},6m}$. That is:

$\mathbf{risk}_{\mathbf{Jones},Stroke}$ **object_of** $\mathbf{risk\_estimate}_{\mathbf{Jones},6m,Stroke}$

$H\text{-}P_{\mathbf{Jones},6m}$ **object_of** value $\mathbf{risk\_estimate}_{\mathbf{Jones},6m,Stroke}$

Similarly, $\mathbf{risk\ estimate}_{\mathbf{Jones},12m,Stroke}$ is about both $\mathbf{risk}_{\mathbf{Jones},Stroke}$ and $History\text{-}parts_{\mathbf{Jones},12m}$:

$\mathbf{risk}_{\mathbf{Jones},Stroke}$ **object_of** $\mathbf{risk\_estimate}_{\mathbf{Jones},12m,Stroke}$

$H\text{-}P_{\mathbf{Jones},12m}$ **object_of** value $\mathbf{risk\_estimate}_{\mathbf{Jones},12m,Stroke}$

This ontological framework is more economical than the former one, as it introduces only one risk of stroke for Jones (rather than one for each time-length).

Note that in PARTHOOD, the relations between $\mathbf{risk}_{\mathbf{Jones},Stroke}$, $Stroke_{\mathbf{Jones}}$ and $H\text{-}P_{\mathbf{Jones}}$ are different from the relations that hold between those entities in ONLY. For example, in ONLY, $\mathbf{risk}_{\mathbf{Jones},Stroke}$ would be realized only by instances of $Stroke_{\mathbf{Jones}}$; whereas in PARTHOOD, $\mathbf{risk}_{\mathbf{Jones},Stroke}$ would be realized by all processes that are parts of a maximal realization, and the latter always has as part an instance of $Stroke_{\mathbf{Jones}}$.

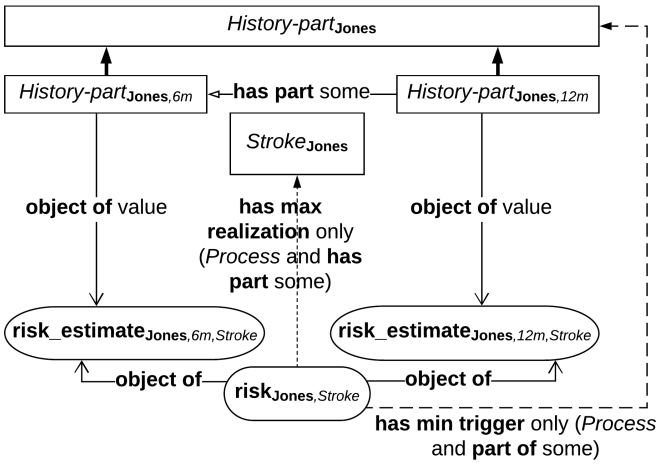

Fig. 4. A formalization of risk of stroke and estimates in PARTHOOD

V. PREDISPOSITIONS

A. Predisposition, risk and disease

In OGMS, a predisposition to disease of type X is defined as "a disposition in an organism that constitutes an increased risk of the organism's subsequently developing the disease X." This definition introduces three dispositions: the predisposition itself, the risk of the organism to develop disease X, and the disease X. Which of those are identical?

First, a predisposition to X is not identical with a disease X: one can have a predisposition to X without ever having a disease X. For the same reason, a risk to X is not identical with a disease X. But what is this relation of "constitution" between a predisposition and a risk? To answer this, we will analyze a specific example.

Consider Mrs. Roy, who has an increased risk of breast cancer because of a BRCA1 mutation. We will distinguish here a causal risk factor from a general predisposition. First, one may say that Mrs. Roy has a causal risk factor to breast cancer because of her BRCA1 mutation. However, despite this mutation, she might have other features (maybe other genetic mutations) that protect her against breast cancer, and therefore she might have overall a lower risk of breast cancer than the average women of the same age group. In such a case, one would not say that she has a general predisposition to breast cancer, taking into account all her features. However, if she has overall a higher risk of breast cancer than other women of the same age group, then she has such a general predisposition to breast cancer relatively to those women. We will analyze in turn the nature of her general predisposition and of her risk factor.

B. General predisposition

Let's consider first Mrs. Roy general predisposition to breast cancer, and show that it can be identified with the risk of breast cancer for Mrs. Roy noted $\mathbf{risk}_{\mathbf{Roy},BC}$. Following [2], the class of triggers of $\mathbf{risk}_{\mathbf{Roy},BC}$ is $History\text{-}part_{\mathbf{Roy}}$, and its class of realizations is the (defined) class $Breast\_cancer\_development_{\mathbf{Roy}}$ (abbreviated in the following as "$BC\text{-}dev_{\mathbf{Roy}}$") of Mrs. Roy developing a breast cancer (note that this class might be empty if Mrs. Roy never get breast cancer; see [3] for a discussion of the nature of empty classes, drawing on [9]). Also, the categorical basis $\mathbf{cat}_{\mathbf{Roy},BC}$ of $\mathbf{risk}_{\mathbf{Roy},BC}$ includes her quality $\mathbf{BRCA1}_{\mathbf{Roy}}$ of having a BRCA1 mutation. That is, $\mathbf{risk}_{\mathbf{Roy},BC}$ is the disposition of Mrs. Roy to develop breast cancer, due to the causal interaction between all features of Mrs. Roy (including her BRCA1 mutation) and the processes happening in her body. But having as categorical basis $\mathbf{cat}_{\mathbf{Roy},BC}$, as class of triggers $History\text{-}part_{\mathbf{Roy}}$ and as class of realizations $BC\text{-}dev_{\mathbf{Roy}}$ also seems to capture the intuition behind her general predisposition to breast cancer $\mathbf{predisposition}_{\mathbf{Roy},BC}$. By application of the criterion of identity ID, since $\mathbf{predisposition}_{\mathbf{Roy},BC}$ and $\mathbf{risk}_{\mathbf{Roy},BC}$ have the same classes of triggers and of realizations, as well as the same categorical basis, they are identical.

More generally, consider $Pred_{BRCA1,BC}$ the class of predispositions of women with the genetic mutation BRCA1 to get a breast cancer. By the same reasoning as above, it can be identified with $Risk_{BRCA1,BC}$, the class of risks of women with the genetic mutation BRCA1 to get a breast cancer.

But when is a risk a predisposition? Not all risks are predispositions: for example, people with drepanocytosis have – like everyone – a risk to get malaria if bitten by a malaria-infected mosquito, but their risk is lower than the risk of people who do not have drepanocytosis; thus, their risk to get malaria cannot be called a "predisposition". And indeed, OGMS suggests that a predisposition is "A disposition in an organism that constitutes an *increased* risk of the organism's subsequently developing the disease *D*." Thus, it is important to capture this

notion of "increased" risk. A natural proposal is to consider that a person **p** has an increased risk to disease *D* if she has a higher probability to get *D* than the probability to get *D* for a class of persons $People_1$ of which **p** is an instance (although this is not the only option: uncertainty might be quantified with other tools than probability).

Probabilities can be analyzed through an epistemic account, as proposed by [2]. In this framework, an epistemic probability *r* can be assigned to a risk estimate that is about a specific risk. Thus, suppose that we have a risk estimate of $\mathbf{risk}_{\mathbf{Roy},BC}$ to which we can assign the probability $p_{\mathbf{Roy},BC}$:

$\mathbf{risk}_{\mathbf{Roy},BC}$ **object_of** some (*Risk_estimate* and **has_value** $p_{\mathbf{Roy},BC}$)

We can compare this risk estimate to the risk estimate assigned to a class $People_1$ of which Mrs. Roy is an instance. For example, if Mrs. Roy is a 54 years-old Quebec woman in 2018, $People_1$ might be the (defined) class of women between 50 and 60 years old in Quebec in 2017. A risk estimate might then estimate a class of risks borne by the class of persons $People_1$:

$Risk_{People1,BC}$ **object_of** some (*Risk_estimate* and **has_value** $p_{People1,BC}$)

Then, we can say that Mrs. Roy has a predisposition to breast cancer if $p_{\mathbf{Roy},BC} > p_{People1,BC}$.

Note that to estimate $\mathbf{risk}_{\mathbf{Roy},BC}$, we would usually also use a risk estimate assigned to a class of which Mrs. Roy is an instance. Suppose for example that Mrs. Roy has not only a mutation BRCA1, but also another genetic protective factor against breast cancer called X. Consider $People_2$ the subclass of $People_1$ who have BRCA1 mutation; and $People_3$ the subclass of $People_2$ who also have the protective factor X. Then, we might provide an estimate of $\mathbf{risk}_{\mathbf{Roy},BC}$ with a probability equal to the probability of a risk estimate of $\mathbf{risk}_{People3,BC}$ (that is, $p_{\mathbf{Roy},BC} = p_{People3,BC}$). This shows that whether a risk is a predisposition or not depends on the choice of a reference class: Mrs. Roy has a predisposition to breast cancer relatively to the class $People_1$ if $p_{\mathbf{Roy},BC} > p_{People1,BC}$, but not relatively to the class $People_2$ if $p_{\mathbf{Roy},BC} < p_{People2,BC}$.

Additionally, even when a reference class such as $People_1$ is chosen, the status of predisposition of a risk depends on the risk estimates that warrants the probabilities such that $p_{\mathbf{Roy},BC} > p_{People1,BC}$. One could imagine a scenario with two different risk estimates of the risk of breast cancer for $People_1$, and two different risk estimates of the probability of breast cancer for Mrs. Roy as follows:

- $\mathbf{risk\_estimate}_{People1,BC}^{1}$ **has_value** $p_{People1,BC}^{1}$
- $\mathbf{risk\_estimate}_{\mathbf{Roy},BC}^{1}$ **has_value** $p_{\mathbf{Roy},BC}^{1}$
- $\mathbf{risk\_estimate}_{People1,BC}^{2}$ **has_value** $p_{People1,BC}^{2}$
- $\mathbf{risk\_estimate}_{\mathbf{Roy},BC}^{2}$ **has_value** $p_{\mathbf{Roy},BC}^{2}$

and such that $p_{People1,BC}^{1} < p_{\mathbf{Roy},BC}^{1}$, $p_{People1,BC}^{2} > p_{\mathbf{Roy},BC}^{2}$: in such a case, Mrs. Roy has a predisposition to breast cancer relatively to $People_1$ and the two first risk estimates, and has no such predisposition relatively to $People_1$ and the two last risk estimates.

For example, $\mathbf{risk\_estimate}_{People1,BC}^{1}$ might be drawn from a first journal article and $\mathbf{risk\_estimate}_{People1,BC}^{2}$ from a second journal article; and $\mathbf{risk\_estimate}_{\mathbf{Roy},BC}^{1}$ might be made by a first doctor who examined Mrs. Roy, and $\mathbf{risk\_estimate}_{\mathbf{Roy},BC}^{2}$ by a second doctor who examined Mrs. Roy. Alternatively, it might be the same doctor who drew two risk estimates $\mathbf{risk\_estimate}_{\mathbf{Roy},BC}^{1}$ and $\mathbf{risk\_estimate}_{\mathbf{Roy},BC}^{2}$, by basing them on two different classes, e.g. basing $\mathbf{risk\_estimate}_{\mathbf{Roy},BC}^{1}$ on $\mathbf{risk\_estimate}_{People3,BC}$ and $\mathbf{risk\_estimate}_{\mathbf{Roy},BC}^{2}$ on $\mathbf{risk\_estimate}_{People2,BC}$ (although the doctor would presumably rely in priority on $\mathbf{risk\_estimate}_{\mathbf{Roy},BC}^{1}$, as it is based on the most specific reference class $People_3$).

When we compare the risk estimate for Mrs. Roy and the risk estimate for a class she is an instance of, both risk estimates need to be made on the same time-frame; that is, both risk estimates need to be about history-parts of the same time-length. Indeed, Mrs. Roy certainly has a higher probability to get breast cancer over 30 years than an unspecified woman would have over one day, but this does not prove that she has a predisposition to breast cancer. It is at least theoretically possible that Mrs. Roy would have a predisposition on a 12-month frame to get breast cancer relatively to the class of women in Quebec between 50 and 60 years old, but would have no such predisposition on a 6-month frame relatively to the same group (if, for example, she has a mild protective factor during the first 6 months that eventually disappear and is replaced by a strong risk factor during the following 6 months).

Note also that not all classes of people would typically be the object of a risk estimate. While a risk estimate might be attributed to the breast cancer risk for the very general class of *Women*, its clinical use would be quite limited given the high variability of individual risks within this class, especially across various ages.

Thus, we can define a predisposition of **o** to *D* as "A risk in an organism **o** of developing a disease *D* about which there is a risk estimate with a probability higher than the probability of a risk estimate on the same time-frame about the risk of developing *D* for a class of organisms of which **o** is an instance."

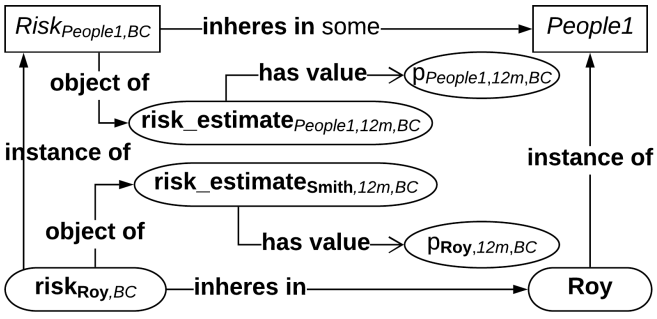

Fig. 5. Individual risk, group risk and their estimates

C. Causal risk factors and protective factors

We can now turn to causal risk factors, as introduced earlier. There is indeed a holistic aspect in $\mathbf{pred}_{\mathbf{Roy},BC}$: it is the general predisposition of Roy to have breast cancer, to which causally contribute all her relevant causal risk factors and causal protective factors (we specify "causal" here, because the

definition of "risk factor" sometimes encompass correlational factor that have no causal influence). However, as mentioned earlier, we might want to represent those risk factors. Suppose now that Mrs. Roy has actually two gene mutations increasing her risk of breast cancer: BRCA1 and BRCA2. Then, we could introduce the risk factor $\mathbf{rf}_{\mathbf{Roy},BC,BRCA1}$ of Mrs. Roy to have breast cancer because of her BRCA1 mutation, as well as her risk factor $\mathbf{rf}_{\mathbf{Roy},BC,BRCA2}$ to have breast cancer because of her BRCA2 mutation. Both are instances of $RF_{\mathbf{Roy},BC,BRCA2}$ such that:

$RF_{\mathbf{Roy},BC,BRCA2}$ **has_trigger** only $History\text{-}part_{\mathbf{Roy}}$

$RF_{\mathbf{Roy},BC,BRCA2}$ **has_realization** only $BC\text{-}dev_{\mathbf{Roy}}$

However, they both have different categorical basis:

- $\mathbf{rf}_{\mathbf{Roy},BC,BRCA1}$ **has_basis** $\mathbf{BRCA1}_{\mathbf{Roy}}$
- $\mathbf{rf}_{\mathbf{Roy},BC,BRCA2}$ **has_basis** $\mathbf{BRCA2}_{\mathbf{Roy}}$

We can quantify the effect of such risk factors as e.g. the difference between the probability assigned to a class with this risk factor and the probability assigned to a class without this risk factor (that is, the "absolute risk increase" [10]). Moreover, we could formalize protective factors in a similar way (and quantify them by "absolute risk reduction"), although they might typically be less used in modern medicine (see the conclusion below for a discussion).

OGMS defines a "genetic predisposition to disease of type *D*" as "A predisposition to disease of type *D* whose physical basis is a constitutional abnormality in an organism's genome." However, this definition rather fits with what we defined as a risk factor. Thus, we suggest to define risk factors and protective factors as follows:

- *Causal risk factor to disease of type* D: "A disposition that causally contributes to increase the probability of a risk of an organism to develop the disease *D*."
- *Causal protective factor to disease of type* D: "A disposition that causally contributes to decrease the probability of a risk of an organism to develop the disease *D*."

To qualify as a reference class for a risk **r**, *R* must be a class whose **r** is an instance (e.g. on Fig. 5, $Risk_{People1,BC}$ for $\mathbf{risk}_{\mathbf{Roy},BC}$).

VI. CONCLUSION

We see that introducing the framework for the identity of dispositions PARTHOOD enables a more economical ontological framework for representing diseases, and suggests a more economical framework for the representation of risks. Moreover, the criterion of identity ID suggests to identify the general predisposition to a disease with the risk to this disease, when this risk is estimated to have a higher probability than the risk associated with a reference class. We have contrasted general predispositions with causal risk factors.

Note that a good understanding of the nature of risk factors, protective factors and general predispositions are of crucial importance for precision medicine. Indeed, a trend sometimes criticized in medicine is to ground prevention or therapy on risk factors, rather than on a general risk level [11]. Moreover, although risk factors are known independently on each other, they synergy is often poorly known. Also, there might be an asymmetry in modern medicine in that it often concentrates on risk factors rather than on protective factors. A good indication of this is that although the term "risk factor" is well-established, the term "protective factor" is much less established. This could lead to an over-medicalization of some naturally protected groups. It is therefore important to have suitable ontological tools to annotate data concerning both kinds of factors.

This analysis also points that the term "predisposition" may have a wider definition than the one presupposed by its regular use. A person is often deemed to be "predisposed" to a disease if she has a risk with a higher probability than the risk of a group to which she belongs of the same age-group and nationality. However, one could imagine to state that a person is "predisposed" to have a disease relatively to, say, younger persons, or persons from a different national group.

Finally, risk factors and protective factors can interact in multiple ways to contribute to a general risk (that might be – or not – a predisposition). Future works should analyze how they can combine in various specific examples to illustrate the large diversity of possible interactions.

ACKNOWLEDGMENT

We thank Ludger Jansen and Alan Ruttenberg for useful discussions on dispositions, as well as anonymous reviewers for their comments. Adrien Barton and Olivier Grenier thank the CIHR-funded "Quebec SPOR Support Unit" for funding.

REFERENCES

[1] R.H. Scheuermann, W. Ceusters, B. Smith, Toward an ontological treatment of disease and diagnosis, in: *Proc. 2009 AMIA Summit Transl. Bioinforma.*, San Francisco CA, 2009, pp. 116–120.

[2] A. Barton, L. Jansen, A. Rosier, J.-F. Ethier, What is a risk? A formal representation of risk of stroke for people with atrial fibrillation, in: *Proc. 8th Int. Conf. Biomed. Ontol.*, Newcastle, UK, 2018, pp. 1–6.

[3] A. Barton, O. Grenier, L. Jansen, J.-F. Ethier, The identity of dispositions, in: *Proc. 10th Int. Conf. Form. Ontol. Inf. Syst. FOIS 2018*, in press, pp. 1–14.

[4] A. Barton, L. Jansen, J.-F. Ethier, A taxonomy of disposition-parthood, in: A. Galton, F. Neuhaus (Eds.), *Proc. Jt. Ontol. Workshop 2017*, CEUR Workshop Proceedings, 2018, pp. 1–10.

[5] M. Horridge, P.F. Patel-Schneider, OWL 2 web ontology language manchester syntax, *W3C Working Group Note* (2009).

[6] B. Motik, P. F. Patel-Schneider, B. Parsia, OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax (Second Edition), 2016.

[7] J. Röhl, L. Jansen, Representing dispositions, *Journal of Biomedical Semantics* **2** (2011), S4.

[8] R. Arp, B. Smith, A.D. Spear, *Building Ontologies with Basic Formal Ontology*, The MIT Press, 2015.

[9] S. Schulz, M. Brochhausen, R. Hoehndorf, Higgs Bosons, Mars missions, and unicorn delusions: How to deal with terms of dubious reference in scientific ontologies, in: *Proc. 2nd Int. Conf. Biomed. Ontol. ICBO 2011*, CEUR Workshop Proceedings, 2011, pp. 183–189.

[10] G. Gigerenzer, W. Gaissmaier, E. Kurz-Milcke, L.M. Schwartz, S. Woloshin, Helping doctors and patients make sense of health statistics, *Psychological Science in the Public Interest* **8** (2007), 53–96.

[11] S. Wells, A. Kerr, S. Eadie, C. Wiltshire, R. Jackson, 'Your Heart Forecast': a new approach for describing and communicating cardiovascular risk?, *Heart* **96** (2010), 708–713.