

## Research Article

# Addiction Ontology: Applying Basic Formal Ontology in the Addiction domain

Janna Hastings<sup>1</sup>, Sharon Cox<sup>1</sup>, Robert West<sup>1</sup>, Caitlin Notley<sup>2</sup>

1. University College London, University of London, United Kingdom; 2. University of East Anglia, United Kingdom

Ontologies are being used in many areas of science to improve clarity and communication of research methods, findings and theories. Many of these ontologies use an upper level ontology called Basic Formal Ontology (BFO) as their frame of reference. This article summarises Basic Formal Ontology and shows how it can provide a basis for development of an Addiction Ontology that encompasses all the things that addiction researchers, practitioners and policy makers want to refer to. BFO makes a fundamental distinction between what it calls continuants (e.g. objects and their characteristics) and occurrents (e.g. processes). Classifying addiction-related entities using this system enables important distinctions to be made that are frequently overlooked or confused in the literature due to inherent ambiguities in natural language expressions. The Addiction Ontology uses this framework to convey information about: people and populations and their characteristics (e.g. substance use disorder), products (e.g. heroin, tobacco-containing products), behaviours (e.g. cigarette smoking, alcohol consumption), interventions (e.g. detoxification, rehabilitation, legislation), research (e.g. measurement, theories, study designs), organisations (e.g. pharmaceutical industry, tobacco companies), and settings (e.g. hospital outpatient clinic, country).

## Introduction

Progress in research and evidence-based treatment of addiction is being hindered by a lack of clarity over the core entities that form the subject matter of the domain. Addiction is a complex and multifaceted phenomenon, and as such many different theoretical and methodological approaches are needed to address its scope in full, arising from such diverse fields as psychology, sociology, medicine, neuroscience and biochemistry. Being able to integrate research arising from within these diverse

fields and perspectives is essential to gaining an overview of what is currently known with what confidence, and specifying where the gaps and conflicts are. However, research synthesis and integration is hampered by the fact that discourse in the field - and accompanying data - often uses terminology without being precise about what the terminology is intended to represent. Thus, the same terms are used to mean different things, different terms are used to mean the same thing or indeed terms may be used incorrectly. This makes extracting information for the purpose of evidence synthesis laborious and inefficient. Examples of entities within the domain of addiction that have been used to mean different things by different authors and across different articles include craving, social norm, poverty, motivation, goal, dependence and identity.

Ontologies are computational logic-based encodings of structured controlled vocabularies that provide unambiguous definitions for entities in a domain as well as formal interrelationships that capture the definitions in a way that computers are able to use. They have been widely used in the biological and medical domains for standardised data annotation to enable integration and reduce ambiguity in widely used terminologies. The Gene Ontology<sup>[1]</sup> was created for the purpose of unifying annotations of gene function across model organism databases and has since grown to become essential to the modern practice of data-driven large-scale genomic science. Many ontologies have since been developed for different aspects of the life sciences domain, including for chemicals, anatomy, disease and biomedical investigations, and many such ontologies are brought together under a common framework in the context of the OBO Foundry<sup>[2]</sup>. This common framework consists partly of computational aspects, such as the use of the standard ontology language OWL and a set of standards for assigning identifiers and metadata. It also consists of a shared common view of the basic divisions in sorts of things in the world that can usefully be distinguished across scientific terminologies, which is implemented as the Basic Formal Ontology<sup>[3][4][5]</sup>.

This article describes how the BFO framework can be applied to entities in the field of addiction research. We first present the theoretical foundations on which BFO is based, and then the core distinctions in types of entity that BFO recognises, with examples of such entities from addiction research. Thereafter, we discuss several examples of ambiguous terms from the field of addiction research, and how their sense can be disambiguated through application of the BFO distinctions.

## Theoretical and methodological foundations for developing ontologies based on BFO

BFO promotes a methodological approach to the development of ontologies, called “ontological realism”<sup>[5]</sup>, that is based on the theoretical and methodological principles of *realism*, *fallibilism*, *perspectivalism* and *adequatism*<sup>[3][4][5]</sup>.

*Realism* is the principle that there is a reality that exists independently of our conceptual and cultural representations. This is, in the broadest sense, to hold it to be true that, for example, there were trees and plants on the planet before the first humans existed or recognised them as such. This independent reality is what our scientific investigations aim to discover more about. However, it should be noted that this principle does not limit the scope of the domain of what can be incorporated to only those entities that have mind-independent existence in the way that trees and plants do. Social entities and psychological entities are also real entities in the relevant sense: for example, a psychological entity such as a feeling of sadness can exist independently of any conceptual and cultural representations of such a feeling. This principle holds it to be the case that scientific findings aim to represent reality as accurately as possible while recognising that such a representation must always be incomplete and derive from a particular perspective. Within the social sciences, a similar approach is *critical realism* which holds that while ‘something’ is there; and that that ‘something’ has an existentially intransitive reality”<sup>[6]</sup>, we can only ever reveal that through the lens of our human interpretation, or through empirical methods, that enable us to agree a shared, falsifiable, consensus of that reality. The remaining principles of fallibilism, perspectivalism and adequatism provide a full framework supporting the application of realism for ontologies in scientific contexts.

Ontologies need to represent, as far as possible, a shared scientific consensus on the nature of entities: the scope of scientific ontologies is ‘settled’ science, rather than science at the fringe or new entities at their first proposal before they have been subjected to the scrutiny of experimental tests. For scientific domains that are in active theoretical flux, it can be hard to know what is settled and what is consensus science.

One might object that much of the field of addiction is contested in this way, not least the definition of the term ‘addiction’ itself, making this a challenging field in which to apply these ontological methods. However, much of the disagreement is around the terminology and related implications, or

technical measurement strategies for specific types of entity, rather than the evidential knowledge base itself. For example, while there may be dispute about whether ‘addiction’ should be defined in terms of a ‘brain disease’ or a ‘social construction’, few who support the latter would argue against the view that there are differences at the neural level between people who experience powerful motivations to engage in a harmful behaviour and people who don’t, and those who support the former would presumably not dispute the observation that there is a social construction giving a label to a condition that a society has identified as problematic. Moreover, journals such as *Addiction* publish studies that follow the scientific method and intend to convey important factual information, hence there is a scientific evidence basis already that would benefit from being organised and catalogued through use of an ontology and accompanying integrative knowledge-based framework.

Scientific consensus serves as a guide to what there is, but not a doctrine about what there is: the principle of *fallibilism* recognises a fundamental truth of progress in science, namely that scientific theory is subject to continual revision in the light of new knowledge.

The principle of *perspectivalism* reminds us that there are a plurality of alternative, equally legitimate perspectives on reality. This should be reassuring in particular for an interdisciplinary domain such as addiction, for which it is important that no one methodological or theoretical approach should dominate.

Finally, in contrast to approaches that suggest what is called ‘eliminative reductionism’, the principle of *adequatism* emphasises that alternative perspectives on reality are not reducible to any single (more basic) view. Thus, psychology does not reduce to biology and chemistry does not reduce to physics. This is not to say that scientific investigations that aim to determine how phenomena at one level of description can be explained in terms of phenomena at another level of description are not important. Such level-crossing research is at the heart of much scientific progress. But the purpose of ontologies is broader than this one aspect of science. To enable data arising from different disciplines to be integrated, and indeed to have a framework beneath which just such explanatory reductions can be consistently classified, it is essential to be able to represent and clearly define entities on each explanatory level in their own right. Furthermore, the principle of *adequatism* also extends to pluralities of perspectives that are not reductionist, such as between different branches of psychology (e.g. social psychology vs. cognitive psychology).

A core element of the ontological approach is to adequately enable the representation of different perspectives by disentangling uses of vocabulary (labels) from contested definitions, and by providing

definitions that highlight the different entities that are in fact being represented when contested terms are used in practice. This is illustrated below with examples drawn from the domain of addiction.

## Basic Formal Ontology: core distinctions

Each of the distinctions in BFO corresponds to a different fundamental type of entity, with different attributes. Figure 1 illustrates the hierarchical (taxonomic) structure of the core categories defined in BFO.

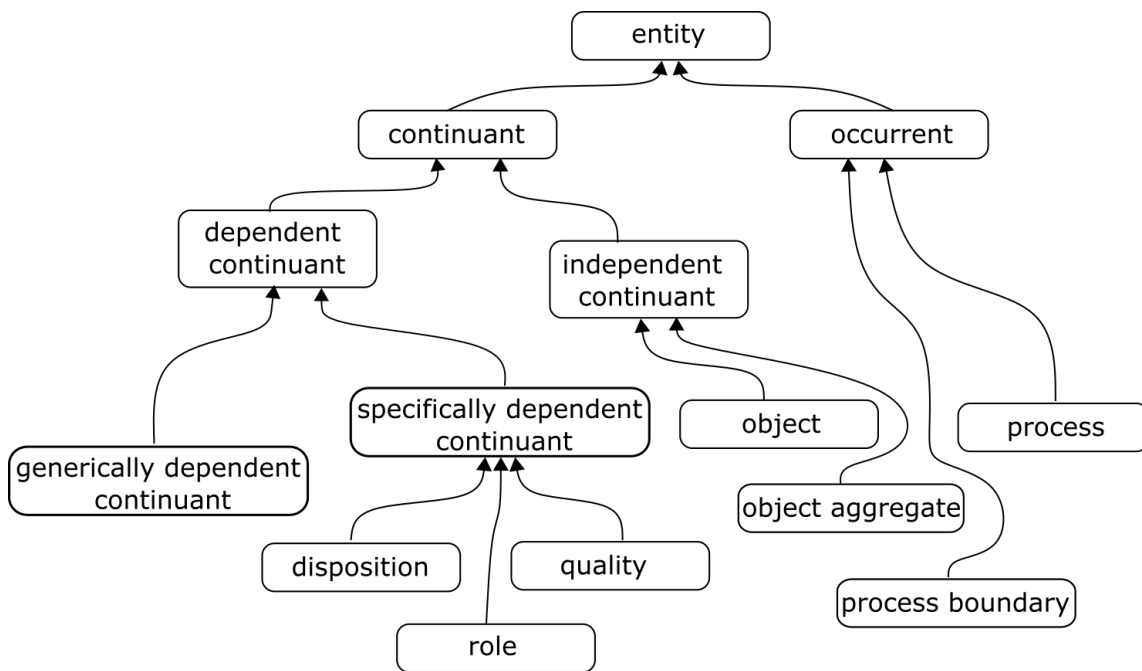


Figure 1: The upper levels of BFO, illustrating the main divisions and types of entity of relevance for addiction.

The most fundamental distinction in BFO relates to the mode of being of entities with respect to time. BFO distinguishes between *continuants*, entities that continue to exist self-identically over a period of time, and *occurrents*, entities that unfold in time and have temporal parts. This is a fundamental distinction that puts, for example, molecules on the one side and chemical reactions on the other; human beings on the one side and conversations on the other. Making this distinction does not deny that continuant entities can gain and lose parts, or change over time, nor does it deny the fundamental wave-particle duality that we know from quantum mechanics is at the heart of the nature of matter

itself. Rather, the distinction emphasises that entities of both of these types form the subject matter of scientific investigations, and therefore both are needed for an adequate description of the subject matter in any given domain.

To determine the mode of being that a particular entity has with respect to time, a helpful thought experiment is to consider whether it exists in full in a momentary 'snapshot' of time – with no movement, no change at all. In such a snapshot one might find a human being, but they would not be breathing, not running or cycling or walking. The human being might be composed of healthy cells, but there would be no metabolism taking place in them. The entities that are still present in a momentary temporal snapshot -- the human, their cells -- are *continuants*, while those that are not -- breathing, metabolism -- are *occurents*.

In the hierarchy of *continuants*, the most important distinction is between what in everyday language we think of as 'things' on the one hand, and the attributes of things on the other. In BFO this distinction is captured by the labels *independent continuant* and *dependent continuant*. Independent continuants are so called because they have inherent existence as objects and regions while dependent continuants only exist as attributes of these. Examples of independent continuants are person, drug, country, brain, and neuronal receptor. Examples of dependent continuants are age, addiction, addictiveness and size.

Note that the dependence that is meant here is not dependence for having come into existence, as a baby depends on its biological parents, nor is it dependence for maintenance of processes that sustain existence, as a human being depends on oxygen. Rather, it is a deeper sort of ontological dependence that conveys the sense that it is *not even conceivable* that any entity of the given dependent type could exist, even momentarily, without a corresponding independent entity in which to inhere. There is no age, for example, that is not the age *of* something. The thing that the dependent continuant is an attribute of is referred to as the *bearer*. Thus when talking about the age of a person the person is the bearer.

BFO makes further distinctions among dependent continuants. One distinction is between what are called *specifically dependent continuants* and *generically dependent continuants*. Specifically dependent continuants are attributes of specific things and if those things cease to exist then so do their attributes. Thus age of a person only exists if the person exists. Generically dependent continuants are abstractions that can be attached to any number of different material things. For example, data exists whether it is stored as ink on paper or electronic patterns in a computer. Study protocols and articles

similarly exist whether as hard copy or stored electronically. Information about things exists as generically dependent continuants.

Then among specifically dependent continuants, BFO distinguishes what it calls *qualities* such as age, which exist in their own right and *dispositions* such as addiction, which only exist by virtue of things that they bring about under certain conditions. Thus the age of a person will always exist if that person exists. People who have 'addiction' only manifest it some of the time.

Dispositions may vary in strength. For example, addiction as a disposition may be stronger or weaker reflecting the frequency and strength of the motivations that are manifested by it.

Dispositions exist because of some attribute or attributes of their material bearers that is non-dispositional in nature. BFO calls this the *material basis* of the disposition. For example, addiction has as its bearer a person and arises out of the characteristics of that person's CNS.

Dispositions and qualities are what ontologists call *internally grounded* in that they are defined completely in terms of their bearers. Addiction is defined in terms of the person who is addicted. By contrast *roles*, e.g. occupational roles such as being a professor, are *externally grounded* in that they are defined in terms of the relationship between the bearer and other entities. Thus being a professor is not intrinsic to a given person but defined in terms of that person's relationship with his or her social environment, including the institution at which they teach and so on. Being an anti-depressant is not something that can be defined purely in terms of a chemical compound; it would be the same chemical even if no-one had ever tried using it in this way. Acting as an antidepressant is a role that the chemical has by virtue of its use. The same chemical, say bupropion, can have the role of anti-depressant and smoking cessation aid.

Turning to the other side of the upper-most distinction, there are also different types of *occurrent*. The most commonly referred to type of occurrent is a *process*, which delineates something happening in time.

Processes always involve at least one kind of continuant as what is known as a *participant*. This term may be confusing to some because in everyday language it usually refers to a person, but in this case it is any continuant that is involved in some way in the process. It could be something that is acting as a causal agent in making the processes happen, something that is undergoing change or something that arises because of the change. For example, the aging process has the person who is aging as a participant. Transmission of a nerve impulse in an axon has the axon and its components as

participants. A research study, which is a process, has human participants in the form of what we would normally think of as such, as well as the researchers and possibly funders, ethics committees and so on.

A very important process in the field of addiction is *history*. This refers to the sum total of all the processes that some continuant, for example a person, has participated in, from their coming into existence to the present time. Entities such as 'ever use' of a drug or past cessation attempts would be defined in terms of this entity.

Picking out boundaries of occurrents is often to some degree arbitrary or specified by *fiat* – a decision made by someone -- and in some cases this is true for continuants as well. Some boundaries, such as the borders between countries, are clearly defined (though they may be disputed). Other boundaries, such as the beginning and ending of adolescence, are to some degree arbitrary. Although it is not part of the BFO vocabulary, entities with these kinds of boundaries can be thought of as *fuzzy*, to use terminology that has been popularised in mathematical contexts. In such cases it is imperative for anyone referring to such an entity to clarify where they are drawing the boundaries. Failing to do is a major source of unnecessary confusion and disagreement in the field of addiction.

## **BFO and the Addiction Ontology**

Addiction is building an Addiction Ontology (AddictO) <sup>[7]</sup> as the basis for a comprehensive knowledge base in the field and to support communication and use of research findings. AddictO is also being linked progressively to a Paper Authoring Tool (PAT) that is designed to radically improve the efficiency and effectiveness of communication of research findings ([www.addictionpat.org](http://www.addictionpat.org)). AddictO is available for download in its current form at <http://addictovocab.org/addicto.owl> and can be browsed online at <http://addictovocab.org/>. In a closely related effort, we are also building an e-Cigarette Ontology (e-Cig-O) <sup>[8]</sup> to cover the domain of research of relevance for e-cigarettes and their use. E-Cig-O is a wholly contained specialisation within AddictO.

AddictO uses BFO as its upper level ontology and this has proved helpful in shining a light on ambiguous and confused use of terms in the field. Table 1 lists commonly used terms in the field and shows how they can be interpreted as fundamentally different things, resulting in unnecessary confusion.



Table 1: Examples of terms used in the field of addiction with meanings that map onto fundamentally different types of entities

Term	Alternative BFO entity classes	Examples of usage in the research literature	Definitions in Addiction Ontology
Addiction	Disposition  Process	<p>Addiction as a disposition to experience strong motivation to engage in a behaviour such as opioid use<sup>[9]</sup>.</p> <p>Addiction as a behaviour such as smoking to which people can become addicted; behaviours are processes<sup>[10]</sup>.</p>	<p>Addiction (disposition): A mental disposition that involves repeated episodes of abnormally high levels of motivation to engage in a behaviour, acquired as a result of engaging in the behaviour, where the behaviour results in risk or occurrence of serious net harm.</p> <p>Addictive behaviour (process): A behaviour to which a substantial proportion of those enacting it are addicted.</p>
Craving	Process  Quality	<p>Craving as a process of neural activation of dopaminergic pathways<sup>[11]</sup>.</p> <p>Craving as a subjective state<sup>[12]</sup>.</p>	<p>Craving (process): A pattern of activation of neural pathways in the brain that is experienced as a strong motivation to engage in a behaviour.</p> <p>Subjective craving (quality): A subjective state of strong motivation to engage in a behaviour.</p>
Substance use cessation	Process  Process boundary	<p>Substance use cessation as a process in which a person who has used a substance over an extended period of time does not use that substance any longer<sup>[13]</sup>.</p> <p>Substance use cessation as a point in time that a person goes from having used a substance for an extended period of time to no longer using that substance<sup>[14]</sup>.</p>	<p>Substance use cessation (process): A behaviour pattern over a period of time in which a person who has used a substance over an extended period of time does not use that substance, intending not to use that substance again.</p> <p>Substance use cessation onset (process boundary): The ending of a pattern of substance use behaviour.</p>
Treatment	Process  Object	<p>Treatment as a process that involves an attempt to prevent,</p>	<p>Treatment (process): A planned process that is intended to prevent, cure or mitigate a health</p>

		cure or mitigate a health condition <sup>[15]</sup> . Treatment as a drug used to treat a health condition <sup>[16]</sup> .	condition. Medication (object): A drug that has a therapeutic role.
--	--	---	--

Because AddictO uses the BFO framework as an upper level ontology beneath which entities of relevance to addiction are classified, and no single entity can be classified as more than one of the types in BFO, the different senses that are conflated in the literature as described in the above table must be disambiguated into distinct entities in AddictO which are then classified appropriately beneath the different BFO branches. A selection of examples of entities in AddictO with their BFO parent classes and their definitions are illustrated alongside.

## Discussion

Ambiguity in language and word meanings is not unique to any one discipline. It is a feature of the flexibility and richness of language and of the wide diversity of perspectives that are represented in a subject area. However, for evidence synthesis it presents significant challenges. Ontologies are designed to provide a more systematic vocabulary for entities in a domain that can be consistently applied to structure research and findings. As an added benefit, they are computable and thus can be used to power intelligent systems that support research. Ontologies can be applied to structure research results retroactively through annotation of published literature reports (manually and also by means of automated natural language processing techniques such as machine learning). They can also be applied directly to data in open databases, and prospectively to the literature through the association of structured metadata at the time of writing reports (e.g. as facilitated by the Addiction Paper Authoring Tool<sup>[17]</sup>).

One of the key objections to an ontology within the discipline of addiction sciences may come from those who position themselves towards the social constructionist end of the epistemological continuum. Those who hold a more contextualist than realist view may consider it impossible to agree on definitions. However, the application of a formal ontology with a clear upper level structure can still be useful even for members of this community for pragmatic reasons. All that is required is that a given ontological entity is operationalised and contextualised appropriately in its context of use such

as a particular research study. This underscores the importance of developing ontologies with participation from all members of a community across different epistemological perspectives in order to actively broker a pragmatic, working level of consensus sufficient to support evidence synthesis in areas that are still contested.

Adoption of a disambiguating upper level for an ontology helps to keep the structure of the ontology clean of some obvious areas of ambiguity and confusion that cut across domains and for which many different examples can be found. This has many practical applications, for example, the hierarchy of the ontology can be used to derive different levels of aggregation of findings dynamically from annotated data, and to browse associated literature or expand queries to find all related papers.

## Conclusion

Defining entities in a domain in clear, consistent and re-usable ways enables integrative interpretation and ultimately a cumulative science. In many cases in scientific reporting, vocabulary elements are ambiguously used and it is only possible to disambiguate the meaning with the context, if at all. This polysemy and vagueness hinders research synthesis. Applying clear distinctions between different senses, based on a shared representation of the different sorts of entity that exist in the world (i.e. Basic Formal Ontology) has the potential to support clear and specific disambiguation at the time of reporting, leading to the potential for systematic research synthesis.

## References

1. <sup>^</sup>Michael Ashburner, Catherine A. Ball, Judith A. Blake, David Botstein, et al. (2000). Gene Ontology: tool for the unification of biology. *Nat Genet*, vol. 25 (1), 25-29. doi:10.1038/75556.
2. <sup>^</sup>Barry Smith, The OBI Consortium, Michael Ashburner, Cornelius Rosse, et al. (2007). The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nat Biotechnol*, vol. 25 (1), 1251-1255. doi:10.1038/nbt1346.
3. <sup>a</sup>, <sup>b</sup>Pierre Grenon, Barry Smith. (2004). SNAP and SPAN: Towards Dynamic Spatial Ontology. *Spatial Cognition & Computation*, vol. 4 (1), 69-104. doi:10.1207/s15427633scco401\_5.
4. <sup>a</sup>, <sup>b</sup>Robert Arp, Barry Smith, Andrew D. Spear. (2015). Building Ontologies with Basic Formal Ontology. MIT Press.

5. <sup>a, b, c</sup>Barry Smith, Werner Ceusters. (2010). Ontological realism: A methodology for coordinated evolution of scientific ontologies. doi:10.3233/a0-2010-0079.
6. <sup>^</sup>Leigh Price, Lee Martin. (2018). Introduction to the special issue: applied critical realism in the social sciences. *Journal of Critical Realism*, vol. 17 (2), 89–96. doi:10.1080/14767430.2018.1468148.
7. <sup>^</sup>Robert West, John Marsden, Janna Hastings. (2019). Addiction Theories and Constructs: a new series. *Addiction*, vol. 114 (6), 955–956. doi:10.1111/add.14554.
8. <sup>^</sup>Sharon Cox, Janna Hastings, Robert West, Caitlin Notley. (2020). The case for development of an E-cigarette Ontology (E-CigO) to improve quality, efficiency and clarity in the conduct and interpretation of research. *Qeios*. doi:10.32388/5YYRPJ.
9. <sup>^</sup>Jette Højsted and Per Sjøgren. (2012). Addiction to opioids in chronic pain patients: A literature review. *European Journal of Pain*, vol. 11 .
10. <sup>^</sup>David Goldman, Gabor Oroszi and Francesca Ducci. (2005). The genetics of addictions: uncovering the genes. *Nature Reviews Genetics*, vol. 6 .
11. <sup>^</sup>Giovanni Addolorato, Lorenzo Leggio, Ludovico Abenavoli, Giovanni Gasbarrini. (2005). Neurobiological and clinical aspects of craving in alcohol addiction: A review. *Addictive Behaviors*, vol. 30 (6), 1209–1224. doi:10.1016/j.addbeh.2004.12.011.
12. <sup>^</sup>Saul Shiffman. (2000). Comments on craving. doi:10.1046/j.1360-0443.95.8s2.6.x.
13. <sup>^</sup>Denise B. Kandel. (1989). Cessation of Illicit Drug Use in Young Adulthood. *Arch Gen Psychiatry*, vol. 46 (2), 109. doi:10.1001/archpsyc.1989.01810020011003.
14. <sup>^</sup>D. R Longo. (2001). A prospective investigation of the impact of smoking bans on tobacco cessation and relapse. doi:10.1136/tc.10.3.267.
15. <sup>^</sup>T Wickizer, C Maynard, A Atherly, M Frederick, et al. (1994). Completion rates of clients discharged from drug and alcohol treatment programs in Washington State. *Am J Public Health*, vol. 84 (2), 215–221. doi:10.2105/ajph.84.2.215.
16. <sup>^</sup>Eugenia Oviedo-Joekes, Suzanne Brissette, David C. Marsh, Pierre Lauzon, et al. (2009). Diacetylmorphine versus Methadone for the Treatment of Opioid Addiction. *N Engl J Med*, vol. 361 (8), 777–786. doi:10.1056/nejmoa0810635.
17. <sup>^</sup>Robert West. (2020). Addiction Paper Authoring Tool (PAT): A Guide. *Qeios*. doi:10.32388/L2KF6W.

## Declarations

**Funding:** This project is funded by Cancer Research UK (C64552/A29896) and the Society for the Study of Addiction.

**Potential competing interests:** JH has no financial or other competing interests to declare. SC has provided expert consultancy to providers of UK life insurance and to the pharmaceutical industry on matters relating to smoking cessation. She also receives salary support from CRUK (C1417/A22962). CN has no financial or other competing interests to declare. RW is an unpaid director of the Unlocking Behaviour Change Community Interest Company.