Reviewer’s report

Title: "OPTImAL": an ontology for patient adherence modeling in physical activity domain

Version: 0 Date: 23 Jul 2018

Reviewer: Amanda Hicks

Reviewer's report:

The project reported in this paper is important and clearly motivated. The supplementary files are excellent and render the science open and transparent. However, the mechanics of the paper itself need to be improved before this reviewer can recommend the work for publication. Many of the sentences are unclear and many words are not used in accordance with their technical meaning which diminishes the quality of the reporting. With careful attention to the details of the reporting, this paper can make a good contribution to the field.

'human lives losses'

'infarction, are'

'ranging up to 56%' This is not a range. Just say 'up to 56%'

'low level of education strength' - Delete 'strength'

'The practice of cardiologists working with patients before, during or after cardiac rehabilitation, shows that the available information that might clarify the situation with the patient or adjust an intervention is accessed through bibliographic databases (e.g., MEDLINE [9]).' - This an intriguing assertion, yet it lacks detail and support. What practices are being referred to? How was this claim established? It needs a reference.

Table 1 does not show results for searching simply for 'adherence ontology'. It is not clear why these are not included since it may provide useful resources for modelling adherence in general that can then be adapted to physical exercise. For example, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4854235/ looks relevant. There are many more relevant articles in PubMed Central that are not described in the literature review.

The paper does not describe the results of the validation. Table 3 is unnecessary since this is not specific to the project but simply describes what the reasoner does.

"The cardiologist assessed the appropriateness of the ontology." This raises many questions. First, what method did the cardiologist use? Does the cardiologist have background in DL and ontologies? If not, the assessment may not indicative of the value of the ontology for inference.
"What factors related to adherence in patients with heart disease who have depression are studied/found?" - This is ungrammatical

"Once the results of the queries were available, the cardiologist assessed if the result was helpful" - "Helpful" is vague. A better method for evaluation is to have a set of answers that are ground truths, and then evaluate the queries for precision and recall.

What does "suitable" mean in the evaluation phase with CRs?

Query 3 in Table 4 does not mention adherence, and is therefore not a suitable query. If this ontology were integrated with another, the extension of the class expression could include things that are unrelated to adherence of physical activity.

Making the literature review available on Zenodo is very helpful. Bravo! Give a little background about the literature review.

The caption for Figure 2 should replace 'superclasses' with 'top-level classes'. Any class that is not a leaf node is a superclass. The manuscript should also be updated accordingly.

"Examining the domain of the ontology, we found that activity behavior is consistently specified with dimension, stage, and duration." It is not clear what the claim for this sentence is. Since the ontology was constructed by authors, it does not seem like this is a discovery unless further work was done with the ontology.

"Further, indicating a combination of a specific patient activity behavior and adherence type and its settings, we implemented classes equivalent to an intersection of interrelated classes." This is unclear.

"We designed an intersection of the class representing activity behavior (Exercise) and an individual of the class" Something is inaccurately stated here. Protege does not support class expressions that are the intersection of a class and an individual.

"We excluded the rest of the properties from being characterized as functional, inverse functional, transitive, symmetric, asymmetric, reflexive, irreflexive." Why?

It would be useful to the reader to have a description of your motivation for reporting metrics of the ontology abstraction and interpretation of these results.

"Classification of individuals in PatientFactor class was held concerning different patient aspects." - This is unclear.

"We divided 320 discovered factors regarding patient demographics, anthropometrics, physiology and physical health, cardiovascular disease and comorbidity, lab tests, symptoms, cognitive and psychological state, environment, social environment and social support, health
behavior, health literacy, lifestyle, healthcare service and insurance, and exercise physiology and exercise settings." - This is unclear

Instead of saying "individuals under", say "instances of".

The individuals shown in Figure 5 are not instances of the classes, but subclasses. They should be modified accordingly. An instance would be "John Doe's change of job on January 1, 2018", not "change of job."

"Classes representing activity behavior settings (i.e., stage, dimension, etc.) and types of relations between factors and adherence are described through equivalence to individuals as shown in Figure 7." Protege does not support describing classes as equivalent to individuals.

"The retrieved results for the factors variously related to adherence in patients with heart disease who have depression are fear of exercise, lack of knowledge about exercise, low mood, low motivation to exercise, negative perception of health, physical restrictions, positive social support, etc. (Query 3)." It seems that some of these will negatively influence adherence (e.g., fear of exercise) while others will positively influence adherence (e.g., positive social support). Does the ontology distinguish these cases? Also, note the that query does not specify what is related to adherence in patients with depression, but what is studied in the depressed population. Something can be studied without being related, so it seems this class expression does not describe the intended class.

"Additionally, the cardiologist mentioned the advantage of the availability of the evidence supporting sources of the results. In practice, if the doctor desires to be certain about the evidence that will be considered, the developed ontology promotes such feature through annotations (see Figure 9)." This is a real strength of the work.

Figure 9 is difficult to read.

"However, we gave equality to similar found expressions." This is unclear. Speak of equivalence rather than equality.

"However, holding experience in the research data analysis, we see the biases of interpreting the data." This is unclear.

The authors might want to consider the OBO Foundry best practices for creating labels in the ontology to facilitate legibility of the labels.

The subclass of relations in "Additional file 5. List of the individuals by classes" are inaccurate. The owl file lists at least some of these as NamedIndividuals, not owl:Classes. These technical details need to be carefully attended to for accurate reporting.

Will the ontology be available for reuse? If so, under what kind of license? Will it be maintained in the future?
Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

No

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

No

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

Yes

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

Not relevant to this manuscript

Quality of written English
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?
If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal