HAL
open science

Proceedings of a Sickle Cell Disease Ontology workshop - Towards the first comprehensive ontology for Sickle Cell Disease

Nicola Mulder, Victoria Nembaware, Adekunle Adekile, Kofi A Anie, Baba Inusa, Biobele Brown, Andrew Campbell, Furahini Chinenere, Catherine Chunda-Liyoka, Vimal K Derebail, et al.

To cite this version:

Nicola Mulder, Victoria Nembaware, Adekunle Adekile, Kofi A Anie, Baba Inusa, et al.. Proceedings of a Sickle Cell Disease Ontology workshop - Towards the first comprehensive ontology for Sickle Cell Disease. Applied & Translational Genomics, 2016, 9, pp.23-29. 10.1016/j.atg.2016.03.005 . pasteur-02012176

HAL Id: pasteur-02012176

https://riip.hal.science/pasteur-02012176v1

Submitted on 8 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Applied & Translational Genomics 9 (2016) 23–29

Contents lists available at ScienceDirect

Applied & Translational Genomics

journal homepage: www.elsevier.com/locate/atg

Proceedings of a Sickle Cell Disease Ontology workshop — Towards the first comprehensive ontology for Sickle Cell Disease

Nicola Mulder [a,1], Victoria Nembaware [a,1], Adekunle Adekile [b], Kofi A. Anie [c], Baba Inusa [d], Biobele Brown [e], Andrew Campbell [f], Furahini Chinenere [g], Catherine Chunda-Liyoka [h,i], Vimal K. Derebail [j], Amy Geard [k], Kais Ghedira [l], Carol M. Hamilton [m], Neil A. Hanchard [n], Melissa Haendel [o], Wayne Huggins [m], Muntaser Ibrahim [p], Simon Jupp [q], Karen Kengne Kamga [r], Jennifer Knight-Madden [s], Philomène Lopez-Sall [t], Mamana Mbiyavanga [a], Deogratias Munube [u], Damian Nirenberg [n], Obiageli Nnodu [v], Solomon Fiifi Ofori-Acquah [w], Kwaku Ohene-Frempong [x], Kenneth Babu Opap [a], Sumir Panji [a], Miriam Park [y], Gift Pule [k], Charmaine Royal [z], Raphael Sangeda [g], Bamidele Tayo [aa], Marsha Treadwell [ab], Léon Tshilolo [ac], Ambroise Wonkam [k,2]

[a] H3ABioNet Consortium, Computational Biology Group, Department of Integrative Biomedical Sciences, Institute of Infectious Disease and Molecular Medicine, Faculty of Health Sciences, University of Cape Town, South Africa
[b] Department of Pediatrics, Faculty of Medicine, Kuwait University, Kuwait City, Kuwait
[c] London North West Healthcare NHS Trust & Imperial College London, London, United Kingdom
[d] Evelina Children's Hospital, Guy's and St Thomas NHS Trust, London, United Kingdom
[e] Department of Paediatrics, College of Medicine, University of Ibadan, Ibadan, Nigeria
[f] Pediatric Hematology/Oncology and Center for Human Growth and Development, University of Michigan, Ann Arbor, MI, United States
[g] Muhimbili Hematology Program, Dar es Salaam, Tanzania
[h] University Teaching Hospital (UTH), Lusaka, Zambia
[i] University of Zambia (UNZA) School of medicine, Lusaka, Zambia
[j] Division of Nephrology and Hypertension, Department of Medicine, UNC Kidney Center, University of North Carolina at Chapel Hill, NC, United States
[k] Division of Human Genetics, Department of Clinical Laboratory Sciences, National Health Laboratory Service and University of Cape Town, 7925, South Africa
[l] Université de Tunis El Manar, Institut Pasteur de Tunis, LR11IPT06 Laboratory of medical parasitology, biotechnologies and biomolecules, Group of Bioinformatics and mathematical modeling, Tunis, Tunisia
[m] Research Computing Division, RTI International, Research Triangle Park, NC, United States
[n] Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, United States
[o] Oregon Health and Science University, Portland, OR, United States
[p] Institute of Endemic Diseases, — University of Khartoum, Sudan
[q] European Bioinformatics Institute, London, United Kingdom
[r] University of Yaounde, Douala, Cameroon
[s] Sickle Cell Unit, TMRI, University of the West Indies, Jamaica
[t] Department of Pharmacy, Biochemistry Unit, , Cheikh Anta Diop University, Dakar, Senegal
[u] Department of Paediatrics and Child Health, College of Health Sciences, Makerere University/Mulago Hospital, Kampala, Uganda
[v] Centre of Excellence for Sickle Cell Disease Research and Training, University of Abuja, Abuja, Nigeria
[w] Center for Translational and International Hematology, Vascular Medicine Institute, University of Pittsburgh, Pittsburgh, PA, United States
[x] Children's Hospital of Philadelphia, Philadelphia, United States
[y] Instituto da Criança, Hospital das Clínicas, São Paulo Medical School, University of São Paulo, Brazil
[z] Duke University, Durham, NC, United States
[aa] Loyola University Chicago, Chicago, IL, United States
[ab] UCSF Benioff Children's Hospital Oakland, Oakland, CA, United States
[ac] CEFA/MONKOLE, Kinshasa, Congo

ARTICLE INFO

Article history:
Received 1 March 2016
Received in revised form 11 March 2016
Accepted 11 March 2016

ABSTRACT

Sickle cell disease (SCD) is a debilitating single gene disorder caused by a single point mutation that results in physical deformation (i.e. sickling) of erythrocytes at reduced oxygen tensions. Up to 75% of SCD in newborns world-wide occurs in sub-Saharan Africa, where neonatal and childhood mortality from sickle cell related complications is high. While SCD research across the globe is tackling the disease on multiple fronts, advances have yet to significantly impact on the health and quality of life of SCD patients, due to lack of coordination of these

[1] The first 2 authors are joint first author.
[2] The last author is senior author.

http://dx.doi.org/10.1016/j.atg.2016.03.005
2212-0661/© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

disparate efforts. Ensuring data across studies is directly comparable through standardization is a necessary step towards realizing this goal. Such a standardization requires the development and implementation of a disease-specific ontology for SCD that is applicable globally. Ontology development is best achieved by bringing together experts in the domain to contribute their knowledge.
The SCD community and H3ABioNet members joined forces at a recent SCD Ontology workshop to develop an ontology covering aspects of SCD under the classes: phenotype, diagnostics, therapeutics, quality of life, disease modifiers and disease stage. The aim of the workshop was for participants to contribute their expertise to development of the structure and contents of the SCD ontology. Here we describe the proceedings of the Sickle Cell Disease Ontology Workshop held in Cape Town South Africa in February 2016 and its outcomes. The objective of the workshop was to bring together experts in SCD from around the world to contribute their expertise to the development of various aspects of the SCD ontology.
© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Background and motivation

Sickle cell disease (SCD) is the most common single gene disorder in the world. It is caused by a single point mutation (Gluβ6Val) that promotes polymerization of hemoglobin (Hb) S and physical deformation (i.e. sickling) of erythrocytes at reduced oxygen tensions. SCD manifests in multiple phenotypes, including inflammation, hemolysis, micro- and macrovascular obstruction and organ damage. Multiple genetic and environmental factors influence the pathophysiological aspects of SCD that contribute to a highly variable clinical expression in individual patients. It is estimated that 305,800 babies are born worldwide annually with homozygous SCD (SCD-SS), with nearly 75% of hemoglobin SS births occurring in sub-Saharan Africa (SSA) (Piel et al., 2013). Despite this high incidence, life-saving public health programs have not been implemented in most SSA countries, often due to limited health care resources and infrastructures. As a consequence, neonatal and childhood mortality due to sickle cell related complications remains high, and estimates suggest that without intervention, up to 90% of affected children in SSA die by five years of age from SCD (Grosse et al., 2011; Fleming et al., 1979).

Despite numerous SCD research efforts across the globe, integration and coordination of these efforts are lacking and a measureable impact on patient care has yet to be realized. In sharp contrast to SSA, comprehensive clinical care programs in high-income nations like the US have reduced SCD-related early childhood deaths by 70% (Yanni et al., 2009; Vichinsky, 1991). Nonetheless, adults with SCD in high-income countries continue to die at a high rate as a consequence of additional debilitating complications (Chaturvedi and DeBaun, 2016, Hamideh and Alvarez, 2013, McClellan et al., 2012; Shankar et al., 2005) and have poor quality of life (Keller et al., 2014; Ameringer et al., 2014; Anie et al., 2012, Dampier et al., 2011).

There is a major need for research on large cohorts of comprehensively and consistently phenotyped SCD patients to help develop effective therapies across the life span for SCD patients in all parts of the world, particularly in Africa. Despite the high SCD prevalence in SSA nations, most of these regions are short of the resources, infrastructure and capacity required to perform epidemiological, translational, and clinical research. Nonetheless, research on an adequately phenotyped large cohort of SCD patients in Africa is necessary to achieve this goal. This, together with establishing public health care infrastructures, and an appropriate management strategy, has the potential to reduce childhood mortality and improve patient quality of life.

Coordination of research activities on different patient cohorts requires development of a Sickle Cell Ontology that provides a controlled and consistent vocabulary of various definitions of SCD in terms of clinical events, genetic and environmental modifiers, co-morbidities, therapeutics, psychosocial burden and quality of life. Such an ontology, ultimately linked to standard protocols for assessing phenotypes and interventions, will inform the standardized collection of data in large cohort studies, as well as surveillance systems which phenotype and follow up patients and families affected by SCD.

Previous work in this area has focused on standardization of SCD complications or phenotypes (Ballas et al., 2010), disregarding other aspects pertinent to this domain such as diagnostics. These attempts suffer from several limitations, which were pointed out by an independent commentary (DeBaun, 2010) and include lack of a reliable method to reach consensus agreements on terms and their descriptions, and imprecise definitions which failed to differentiate between similar but distinct phenotypes.The Sickle Pan African Network (SPAN) and H3ABioNet, a Pan African Bioinformatics Network for H3Africa (H3ABioNet: Mulder et al., 2015) joined forces to lead the development of an improved SCD Ontology, which aims to address the informational gap in the Sickle Cell Disease field. Specifically it aims to:

- establish community standardized SCD terms and descriptions;
- establish canonical and hierarchical representation of knowledge on SCD;
- work collaboratively with PhenX (consensus measures for Phenotypes and eXposures: Hamilton et al., 2011) to establish standard protocols for assessing SCD phenotypes in limited resource populations.

In this paper, we describe the development of the ontology, culminating in an ontology development workshop. The workshop was truly international, including participants from the United Kingdom, United States of America, Brazil, Jamaica, Kuwait and eleven African, countries. Most of the workshop participants were invited based on their SPAN membership, and the rest of the international delegation was invited based on their knowledge in ontologies or SCD and involvement in SCD consortia, networks and clinics. Of those invited, thirty six were able to attend, and included expertise in SCD, ontologies and bioinformatics. Workshop outcomes were measured by changes in the state of the ontology before and after the workshop. In addition, the workshop was evaluated by administering a survey for participants. A set of recommendations for developing a disease-specific ontology were developed.

2. Ontology development process

To spearhead the creation of the ontology, a working group was established via collaboration between SCD experts (mostly from the SPAN) and H3ABioNet members. A Harmonization Center (including curators and developers) was created to facilitate the ontology development, and a panel of ontology experts was established to ensure that the ontology was developed based on established good practices. The ontology expert panel consisted of members from the European Bioinformatics Institute (EBI) Ontology Group and the Open Biological and Biomedical Ontologies (OBO) Foundry. An iterative ontology development process (illustrated in Fig. 1) was implemented, involving stakeholder engagement and several community based activities facilitated by the working group. Engagement with stakeholders involved

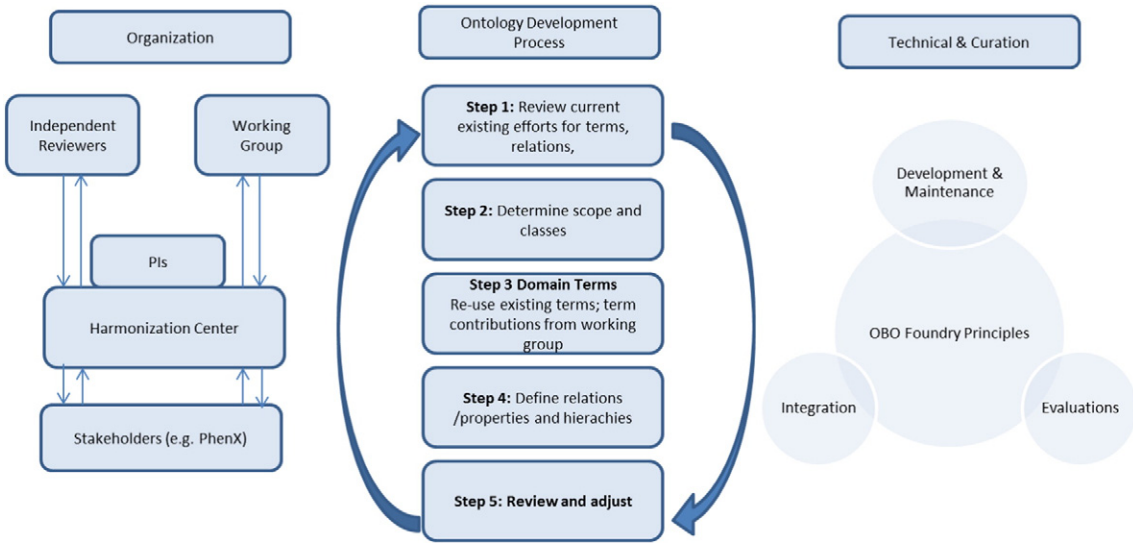

Fig. 1. Overview of the development process.
(Process adapted from Noy & McGuiness, 2001.)

discussions with PhenX project participants to obtain specific terms associated with measures in the PhenX Toolkit (https://www.phenxtoolkit.org/). The EBI contributed phenotype terms cataloged from published text mining tools run against a set of popular SCD-related journals, and H3ABioNet set up WebProtégé (Horridge et al., 2014) for editing the ontology.

Community based activities included meetings with SCD experts; a face-to-face project initiation meeting was held in Livingstone, Zambia, on 11th May 2015, followed by continued online consultations and meetings. A second formal meeting was held during the 7th H3-Africa consortium meeting in Washington, DC, where the draft ontology was discussed, and most recently, an ontology development workshop was held in Cape Town.

3. Workshop proceedings

In February 2016, the Sickle Cell Disease Ontology (SCDO) working group hosted a SCDO workshop. The objective of the workshop was to bring together experts in SCD from around the world to contribute their expertise to the development of various aspects of the SCD ontology. The meeting was attended by 36 participants, including clinicians, geneticists, psychologists and bioinformatics researchers. Sixteen countries, most of which have a high prevalence of SCD, had at least one representative at the workshop. The first day of the workshop was intended to provide background to the participants on ontologies and SCD.

The meeting commenced with a keynote presentation by Professor Kwaku Ohene-Frempong, who gave an overview of SCD and how it is perceived and named traditionally in different parts of Africa. The history of Sickle Cell Disease was also highlighted in this keynote, from the discovery of sickle cells by Dr. Ernest Irons on patient Walter Clement Noel in December 1904 (Herrick, 1910; Savitt, 2010) to the coining of the term "Sickle" by Von M. Lowit in 1905 (Löwit, 1905). The lack of standardized terms in this domain and confusion between the disease and the trait were highlighted. Professor Ohene-Frempong further described how the sickling test may cause confusion around the disease and demonstrated how some medical professionals are not sufficiently educated to effectively treat patients affected with SCD. The keynote was followed by talks introducing the concept of ontologies, how they are created, best practices, and some example ontologies used in the biomedical domain.

The scope of the ontology was based on a draft ontology and on potential use case questions developed via group work. The draft SCD ontology, which had been developed by the Harmonization Center and reviewed by the SCD Ontology working group during H3Africa consortium meetings in Zambia and in Washington, was presented to the participants at this meeting.

Workshop participants were allocated into groups to work on specific classes and sub-classes based on their expertise using the draft ontology as a starting point. Chairs and rapporteurs were allocated to each group.

Once the workshop participants understood the value of creating an ontology, working groups began by defining use case questions. Some example questions were:

- What are the procedures and instruments used for diagnosing patients with SCD with pulmonary hypertension?
- What are the possible adverse effects of treating a pediatric patient with hydroxyurea?
- What is vitreous hemorrhage, and how is it diagnosed and managed?
- What phenotypes are observed when a patient co-inherits sickle cell anemia (SCA) and beta thalassemia?

The second and third days of the workshop focused on group-work aimed at defining the terms and reviewing the sub-classes. The ontology was developed under the following main classes: phenotype, diagnostics, therapeutics, quality of life, disease modifiers (genetic and environmental), and disease stage. Several sources of terms were used for the ontology, these include: SCD experts (electronically and at the workshop), text mining of abstracts using Whatizit (Rebholz-Schuhmann et al., 2008), existing PhenX terms for SCD, Medical Subjects Headings (MeSH), the National Institute of Health and Care Excellence (NICE), World Health Organization (WHO) reports, the Patient Reported Outcomes Measurements Information System (PROMIS) domains (Cella et al., 2010), scientific literature and mining existing ontologies using the EMBL-EBI Ontology Lookup Service (OLS) (Côté et al., 2010). WebProtégé was the tool of choice to collaboratively develop the SCDO and was prepopulated with some of the existing terms curated to date. WebProtégé is an open-source Web application for designing and editing ontologies, which offers several features which permit collaborative development and editing within communities, and has a simple user-friendly interface (Horridge et al., 2014). Some groups chose to continue working in WebProtégé, others used spreadsheets, alternative electronic tools or flip charts.

The workshop concluded with a presentation of the results achieved by each group, followed by a discussion of possible applications of the ontology, and future plans and projects for both the ontology and the working group. An existing SCD database was presented, along with REDCap (Research Electronic Data capture) as an option for an integrated SCD resource. REDCap (Harris et al., 2009) is a secure, web-based application for managing online databases and surveys. It provides multiple sites real-time access to data, while minimizing the logistical challenges in conducting multi-center collaborative projects. REDCap also offers mechanisms for secure storage, validation and reporting of data.

4. Workshop outputs

4.1. SCD ontology development

The workshop participants were committed to the ontology development and lively discussions took place during group work. The ontology was extended most significantly in the fine-tuning of the structure and adding annotations/features to terms. The major changes in the different classes of the draft ontology captured in WebProtégé before and after the workshop are summarized in Table 1.

The Sickle Cell experts at the workshop made recommendations that affected the structure and the contents of the ontology in the following ways: 1) *Class Renaming:* renaming occurs when a class or a property retains the definition and properties that it had prior to the workshop but the name of the class is changed; 2) *Inclusion of Additional Annotation:* some annotations were added to clarify information that was contained in the properties. For example, the *prefLabel* annotation property was added in some classes to resolve naming conflicts, where clinicians had preferred names that differed from ontology terms; 3) *Gap-filling alterations:* these were made in cases where there was pending data to cover already defined classes and properties. For example, many definitions lacked quality references before the workshop and were tagged with <*add source*> in places where a reference was expected; 4) *Complete Redefinition:* this is when a new sub-class or property was defined within a class that previously existed in the ontology version before the workshop. An example of changes to the Diagnostics class is shown in Fig. 2.

4.1.1. Summary reports from the Groups

4.1.1.1. Group I & Group II – Class: Phenotypes. This group was challenged to define clinical terms commonly used in the clinical manifestations of SCD. 1797 terms were extracted from the EBI text mining procedure. Of the 1797 terms reviewed, 1012 terms were approved by the respective experts. Multiple terms were synonyms of each other and these were combined into the most commonly accepted term, as decided by the working group. Six terms have subsequently been added. These terms were classified into the respective systems outlined previously by the working group members. Hematopoeitic, Urinary, Pulmonary, Neurologic, Cardiovascular, Immune, and Musculoskeletal, were among the 18 systems used as "clinical buckets" for their terms. Some of the definitions referenced previously defined terms in publications including the NHLBI 2014 Sickle Cell Disease Evidence Based Guidelines (Yawn et al., 2014) and the Phenotypic Definitions of Sickle Cell Disease publication from United States Comprehensive Sickle Cell Centers (Ballas et al., 2010).

4.1.1.2. Group III – Class: Modifiers. This group suggested differentiating between genetic and environmental modifiers. In addition, this group was tasked with defining the SCD and sickle cell trait. Prior to the start of the workshop, only 4 terms had been defined in this class, and after the workshop this class had 41 terms.

4.1.1.3. Group IV – Class: Diagnostics. Diagnostic procedures used by SCD experts were captured in this class. Most terms in this class were imported from PhenX and existing biomedical ontologies. Discussions included use of terms such as "assay" versus "test". It was agreed to use preferred terminology and use the community recommended ontology terms as synonyms.

4.1.1.4. Group V – Class: Quality of Life and Quality of Care. This group was tasked with defining and standardizing terms in relation to psychological, socio-economic and health care related factors associated with the burden of SCD on affected individuals and their families. Over 60 terms were added within the Quality of Life (QoL) and Quality of Care (QoC) classes. Cross-referencing with other groups was discussed, for example, the potential contributions of socio-economic factors and stress to disease manifestations. A list of psychological and behavioral therapies was shared with the Therapeutics group. Outstanding work includes documenting standard, valid and reliable generic or disease-specific instruments that could be used to capture QoL and QoC information, including measures that can be found in the PhenX toolkit – Sickle Cell Disease Neurology, Quality of Life, and Health Services Specialty Collection. The group will cross-reference clinical symptomatology that has been cited as impacting QoL using the glossary from the National Heart, Lung and Blood Institute Evidence Based Management of Sickle Cell Disease (http://www.nhlbi.nih.gov/health-pro/guidelines/sickle-cell-disease-guidelines), and will also cross-reference the pain descriptors with the Phenotype group.

4.1.1.5. Group VI – Class: Therapeutics. This group focused on the sub-classes related to therapeutic interventions that are already

Table 1
Summary of major changes of metrics in WebProtégé before and after the workshop.

Metric	Class	Sub-class	Terms
Phenotype	Description was added.	The experts deleted the two sub-classes in the draft (Proxy Phenotypes and Related Complications). A classification system similar to the Disease ontology structures was used instead.	Most terms did not have properties before the workshop.
Modifiers	The name was changed from Genetic Modifiers to Modifiers.	Extra sub-classes were added including Environmental modifiers. The Genetic Modifiers class was relegated to a sub-class. The sub-class structures were changed to have deeper branches.	New terms were added to this class.
Diagnostics	Description was added.	New sub-classes were imported from existing ontologies.	Extra terms were imported from existing biochemical assays.
Quality of life	Quality of life was split into two classes: Quality of life and Quality of care.	New subclasses were added.	Terms were adapted from existing vocabularies and reputable sources (references were included). Standard measures were also included.
Quality of care	A new class was created to cater for quality of care.	Five sub-classes were defined under quality of care.	Terms were adapted from existing vocabularies and reputable sources (references were included).
Therapeutics	Description was added	The Abortive sub-class was removed and a new sub-class of Alternative therapies was added.	Terms were defined under each.

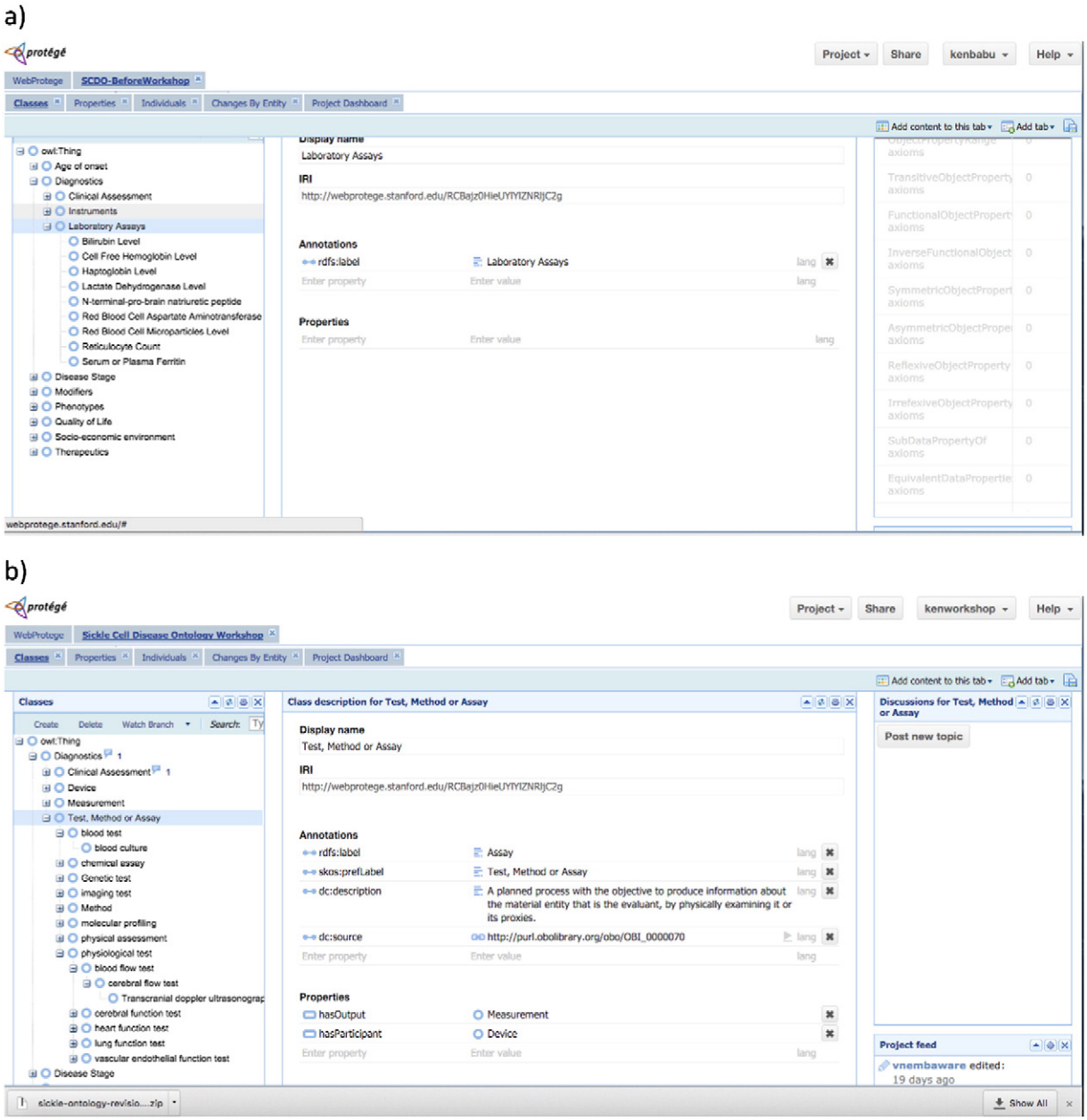

Fig. 2. A snapshot of part of the Diagnostics class from the SCD Ontology in WebProtégé before (a) and after (b) the workshop.

available in WebProtégé. The aim was to identify the different branches that could be associated with each therapeutic class. During the discussion, the group decided to keep the Curative, Preventive, Symptomatic sub-classes, but changed the name of the Supportive sub-class to Supportive/Complementary. They also added a new sub-class named Alternative Therapies and removed the Abortive sub-class. The contents of the Abortive sub-class were moved to the other sub-classes as appropriate.

4.1.1.6. Group VII – Class: Disease stage. Staging of the SCD is still under discussion and is likely to be concluded during the final stages of the development of the ontology. Discussions around staging highlighted the need to differentiate between sporadic, acute, mild, chronic and progressive dysfunction events. Acute events were agreed to be more dangerous, largely unpredictable, and to be the leading cause of mortality among SCD patients. Next steps include reviewing and curating each SCD phenotype with the appropriate disease stage term, where appropriate.

4.2. Workshop recommendations

To finalize the SCD ontology, workshop participants agreed to continue reviewing terms independently after the workshop and make them accessible to their working group for final review. Skype meetings, Google docs and WebProtégé will be used to facilitate the collaborative work going forward. It was also agreed that a face-to-face meeting would be essential to conclude a major section of the work. Once the ontology is finalized, which is expected to be by September 2016, existing SCD data which is comprehensively phenotyped from 2 to 3 different sites in Africa will be used to assess the ontology for coverage and accuracy. The future plan is also to create a standardized case reporting form (CRF) for SCD patient recruitment,

which can be mapped to the ontology and to standard measures (Hendershot et al., 2015).

Based on our experiences and feedback from workshop, we would like to recommend the following plan for building any disease-specific ontology:

- Establish a collaboration between bioinformaticians and disease experts.
- Create a working group to drive the process, and a harmonization center to bring together the ontology work.
- Secure funding for the ontology development and relevant workshops.
- Set milestones with due dates as most contributors are contributing on a voluntary basis. Ensure a mechanism is in place to monitor achievement of milestones.
- Decide on a mechanism for creating, sharing and editing the ontology from multiple sites. Set up a website to host the ontology and related information.
- Organize a workshop (we recommend a full week for such workshops), divide contributors into sub-groups based on their speciality. Translators should be made available.
- Make the workshop and subsequent work as participatory as possible including the planning of the workshop. Set up an emailing list to facilitate communication, continue with ontology development online and provide regular feedback.
- Use real datasets from the community to test the ontology for coverage and accuracy. Encourage the community to start applying the ontology to their datasets and to provide feedback.
- Establish a plan for how to update/modify the ontology as our knowledge base changes over time.
- Have plans in place to disseminate the workshop proceedings in preparation of the releasing of the ontology, through publications and reports. Report on progress regularly to contributors, funders and other stakeholders.

4.3. Workshop evaluation

A survey with four closed and two open questions was administered to workshop participants in order to evaluate the success of the workshop. Twenty-five out of thirty six participants gave anonymous feedback which is summarized in Table 2. Most participants (88%, n = 25) reported intending to use the SCDO. Over 90% of the participants intended to continue being active members of the SCDO working group. Most of the participants recommended longer workshops, better internet connection and translations into French. The participants reported liking the collaborative nature of the workshop, leadership, experts gathered and the productiveness of the workshop. More than half of the participants first learnt about ontologies from the workshop organizers.

5. Conclusions

The SCD Ontology workshop brought together thirty six experts in SCD from all over the world. The enthusiasm and active participation of all attendees was noticeable and infectious. The key challenge will be to maintain this momentum to enable completion and implementation of the ontology. Once completed, we anticipate that the ontology will be the most comprehensive collection of knowledge in the SCD field. It will be used to facilitate exploring of new scientific questions and ideas; to facilitate seamless data sharing and collaborations including meta-analysis within the SCD community; and to support the development and curation of databases and clinical informatics in SCD. Most importantly, it is hoped that the ontology will be used to share data among the SCD community, and enable queries across the different datasets. In addition, we hope that the SCD ontology can serve as a model for other disease communities wishing to establish their own ontology.

Acknowledgments

The workshop was funded by the National Heart, Lung, and Blood Institute (NHLBI), National Institutes of Health (NIH) as a supplement to the H3ABioNet grant. H3ABioNet is supported by the National Human Genome Research Institute (NHGRI), Office of the Director (OD), NIH under award number U41HG006941. The content of this report is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. Funding for two of the participants was provided by the National Human Genome Research Institute and the National Institute on Drug Abuse Genomic Resource Award: U41 HG007050 and the National Heart, Lung, and Blood Institute Supplement: U41HG007050-02S4. We would like to thank Jade Hotchkiss who started the initial work on the ontology.

Table 2
Feedback from 25 workshop participants.

Closed questions	Counts per response
When did you first learn about ontologies?	Prior to workshop — 12 (48%) Prior to the workshop but from the workshop organizers — 9 (36%) During the Sickle Cell Ontology workshop in Cape Town (February 2016) — 4 (16%)
Do you intend to continue being an active member of the Sickle Cell Disease Ontology working group?	Yes — 24 (96%) Not sure — 1 (4%) No — 0
Do you plan on using the Sickle Cell Disease ontology for your own work?	Yes — 22 (88%) Not sure — 3 (12%) No — 0
Please rate the workshop logistics.	No response — 0 Very bad — 0 Bad — 0 Good — 7 (29%) Very good — 17 (71%)
Open ended questions	**Selected quotes**
What did you like the most about the workshop?	"Well organized, expertise well represented, good participation and input from members". "The SCDO workshop enabled me to understand the importance and role of ontology in disease entities in particular SCD"
What changes would you suggest for future ontology development workshops?	"Stronger internet connection so that our work could proceed more efficiently" "I am overall satisfied with the workshops, the only thing I would recommend is include more French specialists for future workshops."

References

Ameringer, S., Elswick, R.K., Smith, W., 2014. Fatigue in adolescents and young adults with sickle cell disease: biological and behavioral correlates and health-related quality of life. J. Pediatr. Oncol. Nurs. 31, 6–17.

Anie, K.A., et al., 2012. Patient self-assessment of hospital pain, mood and health-related quality of life in adults with sickle cell disease. BMJ Open 2.

Ballas, S.K., et al., 2010. Definitions of the phenotypic manifestations of sickle cell disease. Am. J. Hematol. 85, 6–13.

Cella, D., et al., 2010. on behalf of the PROMIS Cooperative Group. Initial item banks and first wave testing of the Patient–Reported Outcomes Measurement Information System (PROMIS) network: 2005–2008. J. Clin. Epidemiol. 63, 1179–1194.

Chaturvedi, S., DeBaun, M.R., 2016. Evolution of sickle cell disease from a life-threatening disease of children to a chronic disease of adults: the last 40 years. Am. J. Hematol. 91, 5–14.

Côté, R., et al., 2010. The Ontology Lookup Service: bigger and better. Nucleic Acids Res. 38, W155–W160.

Dampier, C., et al., 2011. Health-related quality of life in adults with sickle cell disease (SCD): a report from the comprehensive sickle cell centers clinical trial consortium. Am. J. Hematol. 86, 203–205.

DeBaun, M.R., 2010. Finally, a consensus statement on sickle cell disease manifestations: a critical step in improving the medical care and research agenda for individuals with sickle cell disease. Am. J. Hematol. 85, 1–3.

Fleming, A.F., Storey, J., Molineaux, L., Iroko, E.A., Attai, E.D., 1979. Abnormal haemoglobins in the Sudan savanna of Nigeria. I. Prevalence of haemoglobins and relationships between sickle cell trait, malaria and survival. Ann. Trop. Med. Parasitol. 73, 161–172.

Grosse, S.D., et al., 2011. Sickle cell disease in Africa: a neglected cause of early childhood mortality. Am. J. Prev. Med. 41, S398–S405.

Hamideh, D., Alvarez, O., 2013. Sickle cell disease related mortality in the United States (1999–2009). Pediatr. Blood Cancer 60, 1482–1486.

Hamilton, C.M., et al., 2011. The PhenX Toolkit: get the most from your measures. Am. J. Epidemiol. 174, 253–260.

Harris, P.A., et al., 2009. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J. Biomed. Inform. 42, 377–381.

Hendershot, T., Pan, H., Haines, J., Harlan, W.R., Marazita, M.L., McCarty, C.A., Ramos, E.M., Hamilton, C.M., 2015. Using the PhenX toolkit to add standard measures to a study. Curr. Protoc. Hum. Genet. 86, 1.21.1–1.21.17.

Herrick, J.B., 1910. Peculiar elongated and sickle-shaped red blood corpuscles in a case of severe anemia. Arch. Intern. Med. VI, 517.

Horridge, M., et al., 2014. WebProtégé: a collaborative Web-based platform for editing biomedical ontologies. Bioinformatics 30, 2384–2385.

Keller, S.D., Yang, M., Treadwell, M.J., Werner, E.M., Hassell, K.L., 2014. Patient reports of health outcome for adults living with sickle cell disease: development and testing of the ASCQ-Me item banks. Health Qual. Life Outcomes 12, 125.

Löwit, M., 1905. Der Nachweis sichelförmiger Gebilde im myelämischen Blute bei Giemsa-Färbung. Cent. Bakteriol. 39 (39), 274–276.

McClellan, A.C., et al., 2012. High one year mortality in adults with sickle cell disease and end-stage renal disease. Br. J. Haematol. 159, 360–367.

Mulder, N.J., et al., 2015. H3ABioNet, a sustainable Pan-African Bioinformatics Network for Human Heredity and Health in Africa. Genome Res. 26 (gr.196295.115).

Noy, F.M., McGuinness, D.L., 2001. Ontology Development 101: A Guide to Creating Your First Ontology.

Piel, F.B., et al., 2013. Global epidemiology of sickle haemoglobin in neonates: a contemporary geostatistical model-based map and population estimates. Lancet (Lond. Engl.) 381, 142–151.

Rebholz-Schuhmann, D., Arregui, M., Gaudan, S., Kirsch, H., Jimeno, A., 2008. Text processing through Web services: calling Whatizit. Bioinformatics 24, 296–298.

Savitt, T.L., 2010. Tracking down the first recorded sickle cell patient in Western medicine. J. Natl. Med. Assoc. 102, 981–992.

Shankar, S.M., et al., 2005. Medical care utilization and mortality in sickle cell disease: a population-based study. Am. J. Hematol. 80, 262–270.

Vichinsky, E.P., 1991. Comprehensive care in sickle cell disease: its impact on morbidity and mortality. Semin. Hematol. 28, 220–226.

Yanni, E., Grosse, S.D., Yang, Q., Olney, R.S., 2009. Trends in pediatric sickle cell disease-related mortality in the United States, 1983–2002. J. Pediatr. 154, 541–545.

Yawn, B.P., et al., 2014. Management of Sickle Cell Disease. JAMA 312, 1033.