The Standardization of Uveitis Nomenclature (SUN) Project[*]:

Development of a Clinical Evidence Base Utilizing Informatics Tools and Techniques

B. Trusko[1], **J. Thorne**[2], **D. Jabs**[1], **R. Belfort**[3], **A. Dick**[4], **S. Gangaputra**[5], **R. Nussenblatt**[6], **A. Okada**[7], **J. Rosenbaum**[8] **The Standardization of Uveitis Nomenclature (SUN) Project**

[1]Mount Sinai School of Medicine, Ophthalmology, New York, New York, USA

[2]The Johns Hopkins University School of Medicine, Ophthalmology, Baltimore, Maryland, USA

[3]Federal University of Sao Paulo, Department of Ophthalmology, Sao Paulo, Brazil

[4]University of Bristol, Ophthalmology, Bristol, United Kingdom

[5]University of Wisconsin, Ophthalmology, Madison, Wisconsin, USA

[6]The National Institutes of Health, The National Eye Institute, Bethesda, Maryland, USA

[7]Kyorin Eye Center, Tokyo, Japan

[8]Oregon Health and Science University, Ophthalmology and Medicine, Portland, Oregon, USA

Summary:

Background: Given the recent increased focus on evidence-based medicine, it is critical that diseases and syndromes have accurate and complete descriptions, including standardized and widely accepted terminologies. Standardizing these descriptions and terminologies is necessary to develop tools such as computerized data entry forms and classification criteria. This need is especially true for diseases that are relatively uncommon, such as uveitis.

Objectives: To develop a standardized and internationally accepted terminology for the field of uveitis.

Methods: The Standardization of Uveitis Nomenclature (SUN) Working Group (WG) is an international group of 79 uveitis experts from 18 countries and 62 clinical centers. Initial terminology was developed utilizing a "modified" green field approach, which was enhanced through web-based surveys and teleconferences via a "modified" Delphi technique. Terms were mapped provisionally into ontologic dimensions for each syn-drome. The Working Group then met and utilized nominal group techniques as a formalized method of finalizing the mappings. Results: Mapping of terms into dimensions to describe 28 major uveitic diseases was confirmed using nominal group techniques (achieving super-majority consensus) for each of the diseases at a meeting of the entire WG.

Conclusions: The SUN WG utilized an informatics-based approach to develop a standardized and internationally accepted terminology for the uveitides.

Keywords

Classification criteria; uveitis; Delphi technique; nominal group techniques; standardized terminology

1. Introduction

Trends suggest that evidence-based medicine will continue to have a major impact on healthcare in the coming years. The development of automated instruments for the collection of medical evidence, especially related to terminology, classification, and other in-formation contained in the medical record, is key to generating evidence-based medicine on a large scale. This need for a standard language is especially true for relatively uncommon disorders such as those that represent the clinical set of diseases collectively termed uveitis. In relatively uncommon diseases, large databases are not always available, so combined analysis of clinical research from different centers often substitute. Standard terminologies enhance the comparability for this process; however, rigor must be developed which can be used for content discovery and expansion in large, multipurpose vocabularies. Additionally, when utilizing standardized terminologies, we find that uncommon diseases, such as the uveitides, often are not described adequately because they have not been well developed within the scientific/clinical community. These inadequacies in the terminologies for the uveitides are being addressed by the activities of the Standardization of Uveitis Nomenclature (SUN) Working Group.

The development of a quality, statistically significant, robust evidence base is inherently challenging due to several factors, including: 1) with relatively uncommon diseases, a statistically significant sample is difficult to obtain; 2) group meetings tend to elicit descriptions representing the prototypical (textbook) diagnosis (conformity in description), instead of those descriptions being an accurate representation of the actual cases treated in the clinical setting; 3) it is difficult to convene a large, international group of experts in a single room, with access to actual patient medical records and with the time to discuss individual cases (which is the method traditionally used for projects of this type); and 4) coding and analysis of cases from such a diverse representation of clinicians in a timely manner (real-time during a face-to-face meeting) is difficult, especially in the case of a class of diseases with the diverse clinical phenotypes of the uveitides.

Collectively, the uveitides are the 5th or 6th leading cause of blindness in the United States and are responsible for an estimated 30,000 new cases of legal blindness annually, approximately 2.8 –10% of all cases of blindness in the United States. Despite the public health burden, the descriptions of the uveitides often are imprecise [8]. As an example, prior to the 1st SUN Workshop, uveitis experts were polled concerning the use of the term "retinal vasulitis" and given four mutually exclusive options; roughly equal proportions chose each of the options. The consequence is that papers describing risk factors or pathogenetic pathways could be describing different populations and not be com-parable. This problem can be minimized by the development of "classification criteria", which are criteria used for diagnosing specific diseases for reporting them in the scientific literature. The classification

criteria for nearly all of the rheumatic diseases, developed by the American College of Rheumatology, are an example. A key feature of classification criteria is that, although they seek to optimize sensitivity and specificity, when a trade-off is needed, specificity would be favored in order to ensure that a more homogeneous group of patients is being studied. In essence, the greater specificity defines the "phenotype" for clinical research. However, unless basic terms are well defined and widely agreed upon, these criteria's usefulness may be undercut.

In an effort to address these challenges, the SUN Working Group was formed in 2004 to standardize research approaches for the field of uveitis. With the endorsements of the three major uveitis societies, The American Uveitis Society (AUS), The International Uveitis Study Group (IUSG), and the International Intraocular Inflammation Society (IOIS), the first inter-national meeting of the SUN Working Group was held in Baltimore, MD (USA), that year. This was a prototypical working group meeting where experts met over a two-day period and used nominal group techniques to achieve consensus on several generic issues related to research reporting in the field. For example, the 1st SUN Workshop achieved consensus on anatomical site designation, grading schema for anterior segment and vitreous inflammation, and the beginning of a process of developing international standards for reporting clinical data in the field of uveitis, which led to publication of these results. The meeting was successful in advancing several aspects of the nomenclature, and two years following its publication in 2005 it was the most quoted article from the American Journal of Ophthalmology (Liesegang T, personal communication, 2008). An example of the success of this group was the reduction in the number of grading schema used in the literature for anterior chamber cells from six (with variably 6 to 9 steps, depending on the scheme) to one, the number of schema for anterior chamber flare from four to one, and the number of schema for vitreous inflammation from three to one, and a widely- accepted definition of "improved" inflammation, resulting in better comparability of clinical studies reporting "improved inflammation".

In 2009 the SUN WG reconvened to begin development of a complete standardized terminology from which classification criteria could be built. During early discussions, it was clear that a methodology needed to be developed that allowed for: 1) the creation of a set of agreed upon standardized terms capable of describing all the most common uveitic disorders; 2) which could be mapped onto a limited number of ontologic dimensions (e.g. anatomic location of the inflammation, onset, duration, laterality, morphologic appear-ance, etc.); and 3) could be used subsequently to generate the needed classification criteria for each uveitic disorder. Such a process would require the use of informatics methodologies to elicit the group's expertise. Web-based technology and face-to-face meetings would be used to create a terminology of the uveitides sufficient to represent the 28 major clinical entities. Herein we describe the approach the working group utilized for the development of a new highly specific terminology.

2. Methods

From the outset, an "informatics oriented" approach was applied to the problem of obtaining consensus from an "expert" community of international physicians. Taking into account the

difficulties inherent in the development of consensus from a large di-verse international group, which normally is achieved in face-to-face meetings, the Working Group was tasked with creating a new aggregate process to create a terminology acceptable to an international group of uveitis experts. The group was cognizant of the need to convene a "virtual continuous" working group through the use of technology and facilitation that would only be possible through the utilization of management and technology tools. These tools would include a coordinated use of primary and secondary data collection tools, conference calls with working group members, on-line surveys, and live face-to-face meetings. Some of these tools were commercially available and some needed to be developed de novo. Additionally, be-cause of the complexity of the project, electronic tools alone would not be sufficient, and we would be required to utilize both Delphi and nominal group techniques (both virtually and in face-to-face meetings) to reach consensus that was both acceptable to clinicians and would meet the standards required for selection of cases for research by the scientific community. Accordingly, the project consisted of three distinct "phases": 1) collection of primary and secondary data, 2) computer aided, internet-based Delphi technique (DT), and 3) in-person nominal group techniques (NGT).

2.1   Collection of Primary and Secondary Data

During the summer of 2009, core group members (the Executive and Steering Committees (see Acknowledgments)) met to utilize, as a starting point, the recommendations from the 1st SUN Working Group meeting in an effort to develop a working base terminology for approximately 28 uveitic disorders. The team utilized a "modified green field" approach to the development of the initial terms by soliciting descriptions of syndromes from within the working group. This base terminology "considered", but was not bound by, existing terms from SNOMED-CT, LOINC, MeSH and CPT-9, along with the known limitations, (contained in databases in the Mount Sinai School of Medicine Bio-medical Informatics lab), but did not start with them. Therefore, the process was "modified" because we did not attempt to "force" the participants to utilize existing terms; instead we allowed them to utilize descriptions used in practice and uveitis literature. We also utilized the considerable knowledge of the members of the SUN Working Group, academic publications, and other current academic knowledge of uveitis for development of this base terminology. The base terminology was developed after a series of conference calls among two to four members of the team. This was considered essential to the project, since starting with a complete "green field" was considered impractical and inefficient, and offered little to improve the final outcome of the project. Additionally, our initial work discovered that the existing structured terminologies were inadequate (see discussion below) for a complete de-scription of the diseases we were capturing. Finally, the team would use Protege 4.1 Beta (build 213) (http://protege.stan-ford.edu/), to allow us to evaluate relation-ships and create a working model of the project.

2.2   Modified Delphi Technique

The traditional DT is an iterative, multi-stage method that forces consensus among participants without a need for a face-to-face meeting. Participants are given a series of ever-decreasing-in-number forced choices, progressively narrowing the number of options, much like a series of "run-off" elections. It is amenable to Inter-net voting, making it

effective for the inter-national SUN Working Group because conventional communication methods (e.g. conference calls) would be impractical (due to time zone differences, clinical responsibilities, and schedules). Because of the tools available through the Internet, such as weblogs, the SUN Working Group utilized a "modified" DT consisting of three steps.

The modified Delphi approach allowed us to utilize the valuable iterative multi-stage approach of traditional Delphi methods to gather the leading uveitis experts in the world to a scale that had not been done in this way previously. In a traditional Del-phi exercise, several committees are formed and a series of teleconferences take place over some period of time. Given the di-verse global population and the inherent difficulties involved in meetings of large groups of experts the team determined that the "modified DT" was the only acceptable approach. The proxy for the questionnaire, feedback, and additional questionnaire approach utilized in traditional DT was to iterate real time through the blog, which was modified in near real time. This diverged from the traditional process because iterations of the feedback were made available almost immediately; much like a conversation between experts.

Initially, a secure internet-based "blog" was developed to begin to iterate the base terminology among the group. This tool enabled the complete team to read the comments of the group members and con-tribute their own. The blog remained open for two weeks, allowing all members of the working group to comment on the pro-posed base terminology. Members of the SUN Working Group were asked to review the base terminologies to determine if they contained a complete list of the terms that they used to describe the syndromes. This allowed the working group members to create extensive lists of terms that they would commonly use to describe any of the listed disorders. Key terms are listed below in Table 1.

The working group did not intend to modify existing terms (SNOMED, LOINC, MeSH, CPT-9, etc.) since the specific aims of this project were to identify, in the language of the specialist, terms that would as-sist the clinician in describing the uveitic disorder. Accordingly, reconciliation to existing structured terminologies was considered, but the decision was made that reconciliation would be undertaken later in the process. This decision allowed us to leverage the existing standardized terminologies while augmenting where necessary.

Much like a traditional DT approach, feedback was incorporated into the inter-mediate base terminology, which was then included in a web-based survey designed to allow the members of the working group to confirm the base terms by voting on their agreement to the intermediate base terminology and its ability to describe a particular syndrome. For example, a uveitic disorder can be described completely with the taxonomy of history, onset, duration, course, laterality, descriptors, morphology, body site, infection, systemic disease, and findings on laboratory and imaging results. Although not all conditions required a term from each category, all conditions required at least one term from a category. On completion of the intermediate base terminology, proposed mappings of the terms into the taxonomic dimensions were developed for each disease by the core development team and presented during the autumn of 2009 to members of the SUN Working Group via a series of nine conference calls.

A password protected secure site was developed and the results of the DT were posted to a team website for all members of the working group to access in preparation for the 2nd SUN "in person" Workshop.

2.3 Nominal Group Technique

Nominal group techniques (NGT) utilize a highly structured meeting to gather information from relevant experts about a given issue. Initially developed for "brainstorming", it often is used for consensus building. It begins with a series of proposals, developed from earlier work, which are presented to a small group of approximately 10 individuals. These individuals discuss in a formal, time-limited (2 minutes per individual) fashion the proposal with comments from each participant required. A vote is taken, and if a supermajority consensus is achieved (e.g. > 75%), the proposal is passed, and the group moves on to the next proposal. If supermajority consensus is not achieved, the formal discussion is repeated once, the proposal modified as needed, and a second vote taken. If the proposal passes, the group moves on to the next item. If a supermajority consensus is not achieved after two rounds of voting, the item is tabled, and the group moves on to the next proposal. Because the process is structured and time-limited, dominance of the discussion by a single individual or a few individuals is prevented, and input from all participants is elicited. In each small group, a facilitator manages the process; a timer is employed; and voting is anonymous. Although manual methods typically are used for voting (e.g. cards with manual tabulation of results), the SUN Working Group has used electronic voting, which provides anonymous voting and im-mediate feedback, improving efficiency over more traditional methods.

The 2nd SUN Workshop was held in February of 2010 at a two-day conference in Miami Florida, USA. There were 42 SUN Working Group members who were able to attend the meeting. The intermediate base terminology, mapped to the onto-logic dimensions from the Delphi activities were presented in an opening session. Participants then were divided into four groups each focused on diseases grouped by affinity (e.g. the anterior uveitides were addressed in one group, the posterior in another, infectious uveitides in a third, etc.). The participants were assigned to one of the four groups after adjustments were made to minimize regionally weighted or academic "pedigreed" groups. Each group was presented with the proposed mappings of terms into dimensions by the facilitator, and NGT used to determine if there was consensus on the proposed mappings.

On the second day, the results of the small groups were summarized in a Power-Point presentation and presented to the SUN Working Group in a general session. During this session, attendees were allowed to discuss the results of the small groups and recommend modifications to the terms and preliminary classification criteria. The large group, by show of hands, was asked to accept any proposed changes from the small groups, and the large group dis-cussed any items that were not accepted immediately.

3. Results

3.1 Collection of Primary and Secondary Data

After multiple phone and web-based meetings conducted during this phase, the core team developed the base terminology and transferred the findings to Protege 4.1 Beta (build 213) (http://protege.stanford.edu/), allowing us to evaluate relationships and create a working model of the project. The working base terminology was used as a starting point for all other work done by the larger SUN Working Group.

3.2 Modified Delphi Technique

Because we did a significant amount of work establishing a base terminology in the primary and secondary data collection stage, we did not have to start with a completely green field. Instead, the team was able to propose an initial taxonomy and set of terms that were modified through the DT exercise, resulting in the base terms shown in Table 1.

The DT identified 11 categories, 33 sub-categories, 126 terms and 39 sub-terms. Many of the terms were unique terms and some existing terms not previously used in this context. As an example, "orange" exists in SNOMED as a urine color and as a size estimate and "yellow" exists as a color, however, the term "yellow-orange" does not exist in SNOMED but is an important color that only exists in the description of lesions in several of the uveidities. This was the case with many of the categories, sub-categories, terms and sub-terms.

If one were to follow only the "findings" category from Table 1 as an example, the resultant subcategories, terms and sub-terms, would be shown per Figure 1.

The results of the modified DT were compiled and published to a team website for in-person discussions in the NGT phase of the project. Additional terms were no longer accepted and although the project allows for additional terms, none have been proposed.

3.3 In-person SUN Workshop Using Nominal Group Technique

The NGT meeting elicited a significant amount of discussion as well as minor modifications to many of the descriptions. In order to ensure the process functioned properly, the facilitator was required to utilize, and guide, the formalized methodology and to be sure individuals spoke only in turn and for the appropriate time. Also, the facilitator gave guidance about what was being discussed and was required to state the revisions, if there were any, to the group and to provide formal record of the meeting.

The result of the NGT was an agreed upon base terminology (by supermajority) that was voted on by the entire group in a general session of the working group. The final terminology was accepted through group consensus at the conclusion of the large group session. The utility of this terminology was confirmed by achieving super-majority consensus on mappings of the terms into the standard dimensions for each of the 28 major uveitic conditions with no failures. Because we achieved a supermajority consensus on the votes of the mappings and terms, for every one of the 28 diseases under study, the project

did not require any additional analysis of the results. An example of a final terminology for Birdshot Choroiditis is shown in Table 2 below.

4. Discussion

This project addressed, in part, the failure of structured terminologies to address in-adequacies in language for uncommon diseases, such as the uveitides. In the case of uncommon diseases, terms tend to be missing due to the lack of diseases, cases or experts. However, any of these uncommon diseases are important to understand. Even if the team had attempted to utilize natural language processing, for example, the lack of terms to cover the descriptions of the uveidites would have proven too great to simply collect the free text descriptions, as there would have been too many missing terms and differential use of existing terms. It has been pointed out in previous work that use of such tools as natural language processing would only be as good as the terms included in the "dictionary", which if populated with a structured terminology would be deficient for uncommon diseases. Additionally, since much of the documentation from ophthalmology includes symbols and drawings, free text may have proven inadequate for capturing complete cases. This is particularly true in describing shapes and variants of color, both of which are critical for an accurate description of many uveidites. Part of the goal of this project was to find standard descriptions for shapes, sizes and colors.

The opportunity to structure the evidence collection instrument to serve the user is a driving principal in the development of information technology systems. In the case of the SUN Project, this approach also allowed the team to utilize DT and NGT (not commonly used in systems analysis and design) to capture terminologies and mappings through primary and secondary data collection. The team was then able to aggregate the data into a Delphi-ready, internet-based tool and cre-ate terms and mappings consistent with a collection of the worlds foremost experts on uveitis. Through the use of the NGT, we were then able to confirm the DT that can then be used to develop collection instruments to confirm the terminologies. These collection instruments will include a web-based data collection form with "drop-down" menus to collect information on a data set of the uveitides. This data set then can be used for evaluation of proposed classification criteria to develop a parsimonious set of criteria for each disease that optimizes sensitivity and specificity vs. other diseases directly in the differential diagnosis. For example, birdshot chorioretinitis (BSCR) and multifocal choroiditis with panuveitis (MFCPU) both are chronic, bilateral, posterior uveitides but differ in the appearance of the inflamma-ory lesions in the choroid, the amount of inflammation in the anterior chamber and vitreous, and in the mechanisms of visual loss. The clinical inflammatory lesions in BSCR are "yellow-orange" and "ovoid", whereas those of MFCPU are "round" and "punched out and atrophic" in appearance. The development of this terminology permits widely accepted, standardized descriptors and will facilitate future development and evaluation of classification criteria.

5. Conclusions

As the global medical community increasingly shifts focus to evidence-based medicine, one of the greatest challenges is integrating information from diverse sources, a process that will

be facilitated by development of widely accepted standardized terminologies and disease descriptions. The SUN Project was designed to standardize the terms and descriptions of a class of dis-eases that often are described inadequately or are sufficiently uncommon that they are difficult to identify for physicians and re-searchers who do not specialize in that dis-ease.

The approach we utilized allowed the team to evolve from an expert opinion base to a widely accepted base terminology developed by an international team of recognized uveitis exerts. This method would be preferable, for uncommon diseases, to the typical terminology development process. Specifically we were able to perform several iterations of term development within a highly specialized community of experts on a set of linked but relatively uncommon diseases. This allowed us to create a more granular view for each progressive phase of the project. By approaching the problem this way, we were also able to gain consensus with the same group of professionals that will be studying the disease, thereby facilitating research and submit modifications to such groups as the IHE Ophthalmology working group in the future.

A difficulty for future work will be the collection of enough cases of uncommon conditions to convert these preliminary mappings into an optimized and parsimonious set of classification criteria. In -volving a large group of experts will facilitate the collection of a sufficient number of cases. To that end the SUN Working Group was expanded from the original 59 to 79 investigators. The use of a widely accepted terminology and the development of a standardized data collection tool based on this terminology for describing these cases will permit the collection of such data.

Our methodology and informatics tools combined, contribute significantly to reducing the problems of terminology development associated with diseases such as uveitis. Alternatively, this approach can be utilized in reaching and proving expert consensus for almost any disease, but it is particularly effective for uncommon syndromes and diseases where inadequate structured terminologies exist. Additionally, as computers evolve to assist in diagnosis, the ability to apply criteria agreed upon by a global panel of experts will greatly enhance the value of these tools.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Khairaliah M Are the standardization of the uvei-tis nomenclature (SUN) Working Group criteria for codifying the site inflammation appropriate for all uveitis problems? Limitations of the SUN Working Group classification. Ocul Immunol Inflamm 2010; 18 (1): 2–4. [PubMed: 20128640]

2. Cimino JJ. Desiderata for controlled medical vocabularies in the twenty-first century. Methods Inf Med 1998; 37 (4–5): 394–403. [PubMed: 9865037]

3. Dalkey NC. The Delphi Method: an experimental study of group opinion. Rand Corporation mono-graph. RM-5888-PR, 6 1969.

4. Hsu C, Sandford BA. The Delphi Technique: making sense of consensus. Prac Assess, Research and Eval 2007; 12 (10): 1–8.

5. Nussenblatt RB. The natural history of uveitis. Int Ophthalmol 1990; 14: 303–308. [PubMed: 2249907]

6. Brophy S, Calin A. Ankylosing spondylitis: inter-action between genes, joints, age at onset, and disease expression. J Rheumatol 2001; 28: 2151–2154. [PubMed: 11669148]

7. Prieto JF, Dioz E, Gutierrez JM, et al. Pars planitis: epidemiology, treatment, and association with multiple sclerosis. Ocul Immunol Inflamm 2001; 9: 93–102. [PubMed: 11449325]

8. Rosenbaum JT, Holland GN. Uveitis and the tower of Babel. Arch Ophthalmol. 1996; 114: 604–605. [PubMed: 8619773]

9. Jabs DA, Nussenblatt RB, Rosenbaum JT. Stan-dardization of Uveitis Nomenclature for reporting clinical data. Results of the First International Workshop. Am J Ophthalmol 2005; 140 (3): 509–516. [PubMed: 16196117]

10. Rector AL. Clinical terminology: why is it so hard? Methods Inf Med 1999; 38 (4–5): 239–252. [PubMed: 10805008]

11. Raghupathi W, Umar A. Upper-level ontologies for health information systems. Towards an archetype patterns approach. Methods Inf Med 2011; 50 (3): 285–195. [PubMed: 21336420]

12. Stanford Center for Biomedical Informatics Re-search, Protégé. About Protégé. Accessed November 17, 2010 from http://protege.stanford.edu/overview/.

13. Ruperto N, Meiorin S, Mirela S, Ravelli A, Pistorio A, Martini A. Consensus procedures and their role in pediatric rheumatology. Curr Rheumatol Rep 2008; 10 (2): 142–146. [PubMed: 18460270]

14. Stanford Center for Biomedical Informatics Re-search, Protégé. About Protégé. Accessed November 17, 2010 from http://protege.stanford.edu/overview/.

15. Trusko B, Montella D, Jackson JC, Fitzhenry F, Rosenbloom ST, Brown S, Fielstein E, Elkin PL, Kotter K, Tuttle M, Iannelli RJ, Speroff T. Are post traumatic stress disorder mental health terms found in SNOMED-CT medical terminology? J Trauma Stress 2010 11 19.
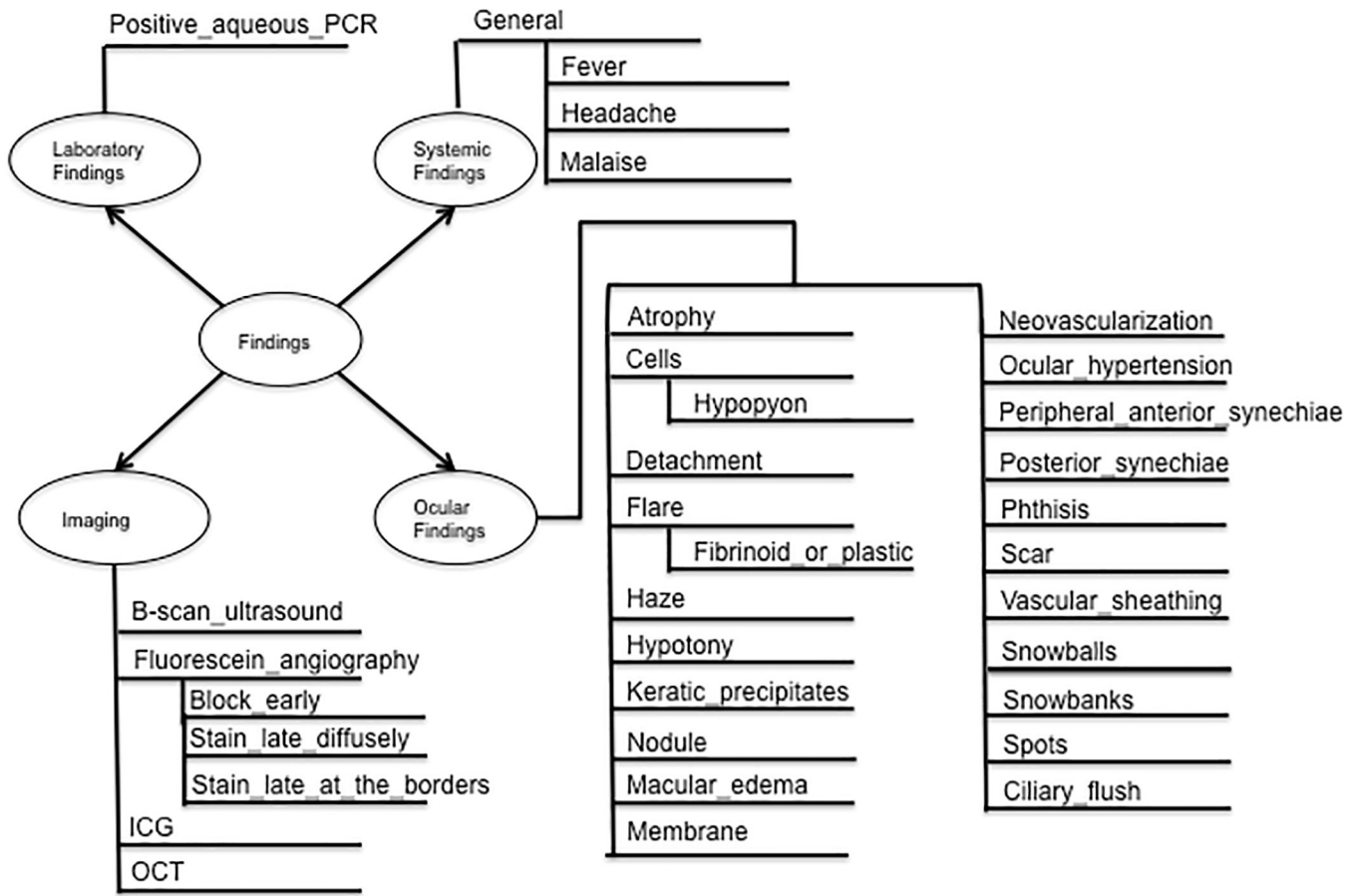
Figure 1:
Findings

Table 1:

Base terms by taxonomy

Dimension	Sub-Category	Terms	Sub-Terms
Body site	1	7	29
Onset	2	0	0
Duration	2	0	0
Course	2	2	0
Laterality	2	6	0
Descriptor	5	59	0
Morphology	3	2	0
Infection	5	10	0
Findings	4	27	8
History	2	0	0
Systemic disease	5	13	2

Table 2:

Birdshot choroiditis mapping

Dimension	Term
Onset	Insidious
Duration	Persistent
Course	Chronic
Laterality	Bilateral, simultaneous
Location	Posterior uveitis
Primary site of inflammation	Choroid
Morphology	Multifocal spots
Descriptors of spots – shape and size	Ovoid indistinct, 50–250 μm
Descriptors of spots – color	Yellow-orange or cream-colored
Fundus location (2-dimensional)	Posterior pole and mid periphery
Other findings	Vitreous cells
Imaging – fluorescein angiogram	Vascular leakage; spots not visible to faint hyperflu orescent