The NIH Office of Rare Diseases Research Patient Registry Standard: A Report from the University of New Mexico's Oculopharyngeal Muscular Dystrophy Patient Registry

Shamsi Daneshvari, PhD, Sarah Youssof, MD, Philip J. Kroth, MD, MS
The University of New Mexico School of Medicine, Albuquerque, NM

Abstract

Patient registries remove barriers to performing research by assembling patient cohorts and data in a systematic, efficient, and proactive manner. Consequently, registries are a valuable strategy for facilitating research and scientific discovery. Registries for rare diseases are arguably even more valuable since there is difficulty in assembling cohorts of adequate size for study. Recently, the NIH Office of Rare Diseases Research created a rare disease registry Standard to facilitate research across multiple registries. We implemented the Standard for the Oculopharyngeal Muscular Dystrophy patient registry created at the University of New Mexico Health Sciences Center. We performed a data element analysis for each Common Data Element defined in the Standard. Problems included the use of previous HL7 versions, non-structured data types, and a recent update to the Standard. Overall, the Standard is an excellent first step toward standardizing patient registries to facilitate work on broader questions and promote novel interdisciplinary collaborations.

Introduction

A rare disease is defined as one that afflicts less than 200,000 people in the United States at any given time.[1] Individually rare diseases do not affect a large proportion of the overall population, however collectively they afflict millions of Americans. The total number of people with a rare disease (25-30 million) is equivalent to the number of people diagnosed with asthma.[2] The geographic dispersion of a relatively small number of patients with rare diseases leads to many unique and complicated issues that are exclusive to these groups. This includes the difficulties in finding providers with the appropriate knowledge of the disease, safe and effective treatments, as well as funding to conduct investigations into the natural history, genetics and pharmacological treatments.[1, 3, 4] Even though a rare disease affects no more than 200,000 individuals many of them share similar symptoms and treatments. Therefore, a critical component of advancing rare disease research centers on the establishment of disease registries and associated data.

The creation of patient registries is a tried and true research methodology for studying patients with a common illness or set of characteristics. Registries collect data on cohorts of patients with a disease, illness, procedure or exposure.[5, 6] As such, they contain information on the participants, their disease, treatments and willingness to participate in future research. Registries can be used to determine the natural history of disease, effectiveness of treatments, monitor safety, improve quality and provide cohorts for future research.[6-8] This is of extreme importance for diseases or conditions that are rare and therefore have afflicted individuals spread out over large regions at low density. Registries perform the time intensive aspect of forming a cohort of subjects and organizing the associated data. This proactively reduces the overall burden for each new research endeavor. The work of assembling the cohort is conducted in a centralized and systematic way that leverages economies of scale to make the cohort's data available to many investigators in perpetuity. Registries for rare diseases are even more valuable as the burden of amassing a cohort for rare disease research is significant. As such, a registry of registries, composed of the aggregated registries of all rare diseases would increase future cohort sizes and allow research not yet before possible.[3, 9, 10]

Registries usually contain information detailing a patient's demographic information, medical history, family history, willingness to participate in research and disease specific information. Patient disease registries help to elucidate patients for future trials, provide quality assessments and improve care for patients.[3, 7] Standards within and between registries would allow aggregation of data as well as possible association with information contained in Electronic Health Records.[11] Standards include the data elements and the terminology standards to encode them.

Multiple terminology standards have been applied to registries, such as those from the International Standards Organizations (ISO), SNOMED-CT, LOINC, RxNORM, and the Office of Management and Budget (OMB) for race and ethnicity.[7, 12, 13] A key issue with registries is a lack of, or continually changing, standards for data collection; hence some organizations have focused primarily on collecting contact information to facilitate recruitment for future research trials.[8]

In response to the particular challenges of studying rare diseases, the NIH established the Office of Rare Diseases Research (ORDR) in 2002. The ORDR serves to fund and encourage research involving rare diseases, establish natural histories, and create a Standard of common elements to facilitate data gathering and sharing[1]. The Global Rare Diseases Patient Registry and Data Repository (GRDR) was launched by the ORDR to enable research over multiple sites and increase participation in clinical trials. The GRDR's goal is to help investigators, patients, and advocacy groups by creating a Standard for a data repository of de-identified aggregate data from all rare diseases registries. The data repository will be available, after completion of a pilot study, to investigators for further research and provide a standardized repository housed at the ORDR.

In 2010, the ORDR convened a workshop, "Advancing Rare Disease Research," for the institutes and agencies participating in the pilot project to compare notes and make recommendations to the developing Standard. The attendees noted that having a standard for the entry of registry data would assist, and in some cases enable, organizations to take part in the process of aggregating rare disease data. Additional recommendations included:

- Standardizing registry questions
- Finding common ground across registries
- Developing a minimum common registry model
- Creating a common repository
- Developing a method for sharing information between registries and the data repository.[14]

A Common Data Elements Steering Committee was created to define the Common Data Elements (CDEs) of the Standard to capture the minimum amount of information needed to best support clinical research on rare diseases. The committee subdivided the information into 10 categories (see Table 1), and within each category, elements were further subdivided into required or optional categories.[15]

Table 1. Categories for the Standard's Common Data Elements.

Categories
• Current contact information
• Socio-demographic
• Diagnosis
• Family history
• Birth and reproductive history
• Anthropometrics
• Patient reported outcome
• Medication and devices
• Clinical research and participation
• Contact/communication preferences

The GRDR developed 10 steps for the eventual deposition of de-identified aggregate data into the master repository. Through this process they aim to populate the master repository from all rare disease registries that conform to the Standard:

1. GRDR determine and publish the Standard's CDEs along with the referenced standards used to define each data element.
2. Registry owners will map their registry's data elements to the Standard's CDEs and determine equivalence or modifications that are needed.
3. Registry owners will submit their data element mapping to the GRDR.

4. The GRDR review and approve the suggested mapping or provide feedback to help the registry fix mapping issues.
5. The registry owners will write any transformation code that is needed to make all of the registry's data into the form defined by the Standard's CDEs.
6. The data from the registry is extracted and transformed.
7. The GRDR loads the registry's data into their holding system.
8. The GRDR assesses the quality of the data.
9. The GRDR determines if the registry's data will be accepted or work with the registry's organization to fix problems identified.
10. The GRDR transfers the data to the repository in perpetuity.[15]

The pilot study began in 2012, with the selection of 24 rare disease organizations -- 12 with existing registries and 12 without. The organizations without a registry are being funded to create their registries using the registry Standard. Organizations with existing registries containing their existing and proprietary data standards will determine how well their existing data maps to the new Standard and specifically how the proprietary data elements map to the Standard's CDEs. The ORDR will use the results of this pilot program to improve the design of the Standard.[4]

To date there have been two versions of the Standard's CDEs released. The second and most recent version released on December 30, 2012, included changes to the number of data elements and to other external data standards referenced in many of the CDEs' definitions. An example of an optional Common Data Element in version 2 is 'Nationality.' This data element documents the country in which a person or their family was previously resident. The response is defined using the external terminology standard ISO (International Organization for Standardization 3166-1-alpha-2 code). The first version of the Standard's CDEs is heavily weighted toward use of HL7 v2.3.1, and the second version includes the use of other terminology standards, such as SNOMED and LOINC.[15]

Background

Oculopharyngeal muscular dystrophy (OPMD) is a late-onset, progressive, hereditary muscle disease that leads to severe inability to swallow, weakness of the eyelid muscles, and limb muscle weakness[16]. There is currently no cure for OPMD; symptom management consists of surgical procedures on the eyelids and pharyngeal muscles to correct ptosis and reduce swallowing difficulties, respectively. While OPMD is a rare disease worldwide, the highest prevalence of OPMD in the Unites States occurs among Hispanic New Mexicans due to a founder effect.[17] Since the year 2000, nearly two hundred Hispanic New Mexicans with OPMD have been seen in the dedicated OPMD clinic at the University of New Mexico. OPMD typically comes to medical attention in the fifth decade of life. Previous studies have determined the point prevalence of signs and symptoms of OPMD in small cross-sectional samples.[18-20] There is a dearth of prospective, longitudinal data on disease progression. Thus, the detailed natural history of OPMD needs elucidation.

Two of the authors (SY, PK) received pilot funding in 2012 to create an OPMD registry. The project was approved by the University of New Mexico's Human Research Protections Office (Protocol number: 12-245). Participants are recruited through the OPMD clinic, word of mouth, and a new website about the registry (http://som.unm.edu/programs/opmd/index.html). Individuals must have OPMD or be related to someone with OPMD, provide consent, and complete a questionnaire to be included in the registry. Participants agree to be contacted yearly to update their contact and health information. Currently the questionnaire consists of 57 questions and 190 recorded elements. The selected elements are from multiple validated and/or field-tested sources (see Table 2). To date, 49 participants with OPMD have enrolled in the registry.

Table 2. Data Elements included in the OPMD registry

Source of data element	Number of data elements
• NIH Office of Rare Diseases Research Standard's CDEs	• V.1: 47, V.2: 42
• Swallowing Quality of Life outcomes tool[21]	• 45
• Neuro-QOL Lower Extremity Function Item Bank	• 8
• PROMIS Physical Function Item Bank	• 3
• International Statistical Classification of Diseases codes	• 6
• Current Procedural Terminology codes	• 6
• U. of Rochester Myotonic Dystrophy and FSHD Registry	• 5
• Rare Diseases Clinical Research Network Contact Registry	• 1
• Sydney Swallow Questionnaire[22]	• 1

Methods

Versions 1 and 2 of the Standard's CDEs reference Health Level 7 (HL7) v2.3.1 unless v2.5 was specifically noted.[15] A list of the 47 OPMD registry fields mapped to version 1 of the Standard's CDEs, and 42 mapped to version 2. An itemized list of differences between the versions was created by matching data element to data element of the Standard's CDEs.

One author (SD) performed a systematic and detailed data element analysis on both versions of the Standard's CDEs and their referenced terminology standards. Each element used within the OPMD registry was mapped to the Standard's CDEs and their referenced terminology standards. Any issues related to the mapping of the OPMD registry's data elements or with the referenced terminology standard were recorded in detail. This process resulted in a list of the OPMD registry field, the Standard's Common Data Element number, the terminology standard referenced and the issues identified as being relevant to the mapping/association.

Results

To differing degrees, both versions of the Standard's CDEs reference Health Level 7 (HL7) v2.3.1. In version 1 of the Standard's CDEs, 37 of the 77 elements reference HL7 (with 2 representing HL7 v2.5). In version 2, only 11 of the 75 elements use HL7 v.2.3.1. With the update of the Standard's CDEs there is also an inclusion of LOINC and SNOMED-CT terminology standards as well as the removal of most HL7 data types and tables. Another change is the alteration of certain elements. For example, date of birth was recorded in version 1 but in version 2, only year of birth is recorded. A distinct and important change between the versions is both the elimination of some and the addition of new CDEs to the Standard. See Table 3 for a complete and detailed summary of the changes between the two versions.

Table 3. Summary of the differences between the Standard's CDEs and their referenced terminology standards, Versions 1 and 2

CDEs Version 1	CDEs Version 2
• 35/77 elements reference HL7 v2.3.1	• 11/75 elements reference HL7 v2.3.1
• 2/77 elements reference HL7 v2.5	• 0 elements reference HL7 v2.5
• 5 HL7 data types used	• 3 HL7 data types used
• HL7 tables referenced	• 1 HL7 table referenced for 1 CDE
• LOINC not referenced	• LOINC referenced
• SNOMED-CT not referenced	• SNOMED-CT referenced
• Date of birth	• Year of birth
• Date of death	• Year of death
• 5 digit zip code	• First 3 digits of zip code

Table 3. Continued

CDEs Version 1	CDEs Version 2
• Contact person is required	• Contact person not included
• Specific date of diagnosis and 1st symptom are required	• Specific date of diagnosis and 1st symptom are not included
• Sex is includes genetic and social sex, both required	• Sex is a single element, required
• HL7 table 0163 for yes/no relationships	• LOINC yes/no/refused/don't know or LOINC yes/no
• HL7 table 0063 for relationships	• LOINC relationships
• OMB race categories	• LOINC extended race
• OMB ethnicity categories	• LOINC ethnicity
• AHRQ-modified for insurance type element	• LOINC insurance
• ACS Questionnaire and US Census for educational attainment element >18	• LOINC level of education
• ACS Questionnaire and US Census for educational attainment element <18	• Unreferenced list
• US Census for family income element	• Not referenced, but identical income groupings
• Age at onset and 1st symptom uses HL7 TQ1 segment	• Age at onset has age in years with specified categories
• Rare disease diagnosis ORDR controlled vocabulary	• SNOMED-CT

Depending on the version implemented, the number of the ORDR Standard's CDEs used in the OPMD registry varies. Forty-seven of the 190 elements in the OPMD registry elements mapped to the Standard's CDEs in version 1. Seven of these mapped elements required modifications of the referenced terminology standards responses. In version 2, the total number of the data elements in the OPMD registry that mapped to the Standard decreased from 47 to 42. Ten of these mapped data elements required modified for the OPMD registry. The most important results of the data element analysis for version 1 are listed in Table 4 and a more detailed list of issues for version 2 in Table 5. A full detailed list for version 2 is available online at http://hdl.handle.net/1928/22765.

Table 4. A Short list of the Standard's Common Data Element analysis for Version 1

Common Data Element	Terminology Standard	Comments
• Full name of patient and contact (first, middle initial, last)	• HL7 ST data type (one for each)	• HL7 data type for name exists: XPN
• Telephone number	• HL7 TN data type	• Data type has been deprecated, current data type: XTN
• Street address, city of patient and contact	• HL7 ST data type (one for each)	• HL7 data type for address exists: XAD
• Sex	• Biological: ISCN 2009 Societal: HL7table 0001	• Individuals will most likely not know their chromosomal sex

Table 5. A Subset of the Standard's Common Data Element analysis for Version 2

Common Data Element	Terminology Standard	Comments
• Full name of patient and contact (first, middle initial, last)	• HL7 ST data type (one for each)	• HL7 data type for name exists: XPN
• Telephone number	• HL7 TN data type	• Data type has been deprecated, current data type: XTN
• Street address, city of patient and contact	• HL7 ST data type (one for each)	• HL7 data type for address exists: XAD
• Sex	• LOINC/NAACCR extended sex	• Corrected previous version by combining into one element

Table 5. Continued

Common Data Element	Terminology Standard	Comments
• Race	• LOINC Extended Race	• Most studies continue to use the OMB categories from the US Census • Categories include general race categories and a further break down of categories within those
• Ethnicity	• LOINC Ethnicity	• Most studies continue to use the OMB categories from the US Census • Referenced responses include general ethnicity categories and a further break down of categories within those
• Contact	• Previously HL7 data types	• Removed in Version 2, data collected by many registries
• Consent	• LOINC	• Reference links to the LOINC code 67791-4, which refers to the specific question "Does the patient take any medications" • Trial code, may change
• Registrar	• LOINC	• Reference links to the LOINC code 67791-4, which refers to the specific question "Does the patient take any medications?" • Trial code, may change
• Record of Self Completion	• LOINC	• Reference links to the LOINC code 63513-6, which refers to the specific question "Are you covered by health insurance or some other kind of health care plan?" • Trial code, may change
• Vital Status	• LOINC	• Reference links to the LOINC code 63513-6, which refers to the specific question "Are you covered by health insurance or some other kind of health care plan?" • Trial code, may change
• Nationality	• ISO 3166 1alpha-2 code	• Link within the Standard's CDEs links to ISO 3166 3-code • Version not noted. Terminology Standard update 2-5 times a year
• Educational Attainment	• LOINC level of education	• Trial code, may change • Modifies terminology standard to omit response option
• Patient Reported Outcome Section	• PROMIS	• PROMIS referenced in LOINC using different response categories

Discussion

The OPMD registry housed at the University of New Mexico implemented the ORDR's Standard. We aim to include data from the OPMD registry in the GRDR de-identified aggregate data from all rare disease registries. To achieve this aim, the authors originally created the OPMD registry questionnaire to conform to the Standard version 1. In doing so, we determined that there were modifications needed, issues with terminology standards, and use of deprecated HL7 data types to our original registry design.

Using version 1 of the Standard, the OPMD registry modified 7 of the 47 CDE terminology standard responses to better conform with the requirements specific to the OPMD domain. The reasons for modifications varied

depending on the Standard CDE. For example, the element 'sex' was divided into two required fields: 'biological sex' and 'societal sex.' We felt that the participants would be confused by the inclusion of genotypes under biological sex. Therefore, we deleted the biological sex CDE and only used the societal sex element. Furthermore, due to the age and background of the participants, the investigators chose to include only male and female with the option of writing in an answer under 'other.' Another notable modification was to the CDE of 'Age when first symptom observed.' The field was broken down into 4 specific diagnostic symptoms, each with a section for recording age. This was decided because patients may not know which symptoms were related to the OPMD diagnosis.

In version 2, the OPMD registry required modification of 10 of the 42 CDEs and referenced terminology standards to better align with the requirements specific to the OPMD domain. The increased number of modified OPMD data elements from version 2 of the Standard is partly due to the creation of the OPMD questionnaire from version 1's CDEs and referenced terminology standards. The OPMD questionnaire will soon be updated to reflect the changes that occurred with version 2. The changes between version 1 and 2 were significant, requiring modifications to much of the data stored in the OPMD registry. Most importantly, there are dramatic changes in the terminology standards referenced in the CDEs and therefore the responses possible on the OPMD registry questionnaire. In other words, there are significant changes to the data gathering process with the implementation of version 2. For example, all but one of the referenced HL7 tables used in version 1 was eliminated and most were replaced by LOINC vocabulary. Date fields relating to birth and death were altered to record only the year rather than the specific date. Zip codes were also altered to store only the first three digits rather than the full number. The date and zip code changes can easily be transformed when OPMD data is uploaded into the ORDR's data repository of rare disease registries. However, the change in terminology standards in elements such as race and ethnicity will require changes to the OPMD registry database. For instance, most investigators record a participant's ethnicity using the Office of Management Budget (OMB) standard which only includes "Hispanic" or "Not Hispanic." The use of the LOINC ethnicity standard specifies 9 categories, most of which are usually subsumed into 'Hispanic' in the OMB standard. As a result, if an investigator is only recording 'Hispanic' in a database, there is no way to map it correctly to the subdivided 9 LOINC categories. This may create a problem with the global registry mapping.

Another significant issue for database designers is the use of LOINC codes only for standardizing question responses (i.e., 'yes/no' and 'yes/no/refused/don't know.') The LOINC code referenced for 'yes/no' responses refers to LOINC code 67791-4 which actually is the 'yes/no' response to the specific question: "Does the patient take any medications?" Likewise, the LOINC code referenced for 'yes/no/refused/don't know' refers to LOINC code 63513-6 which actually is the response for the question: "Are you covered by health insurance or some other kind of health care plan?" Although these LOINC codes are defined for specific questions and responses, version 2 of the Standard uses them only to indicate the form of the response and not the questions themselves. This violates the LOINC standard given that the definitions of these codes are very specific to defined questions and the acceptable form of the answers. According to LOINC version 2.42 these referenced codes are "trial codes" and may change in time.[23]

Other issues specific to version 2 include the use of the *ISO 3166-1 alpha 2-code* (two letter country codes) as a terminology standard. The link within the CDEs spreadsheet provided by the Global Rare Disease Patient Registry and Data Repository is to *ISO 3166-1 alpha 3-code*. More troublesome is that no version of the ISO standard is mentioned. The *ISO 3166-1 alpha 2-code* is updated multiple times per year, so future problems will be created if the version is not specified. For example, in 2010 V1-8 of *ISO 3166-1 alpha 2 code* announced the deletion of 'NL,' which was the 2-code abbreviation for the Netherland Antilles and the replacement terms for the three new countries that were formed (CW, BQ, SX) as well as the three territories added to the Netherlands (NL).[24] As such, when merging databases, without versioning, there may be difficulties when former countries are subdivided.

There are also overarching issues with both version 1 and 2. Both versions reference HL7 v.2.3.1, and use data types that have since been deprecated. With the switch to version 2, the number CDEs that referenced HL7 data types have been significantly reduced, leaving only the string (ST), telephone (TN), and text (TX) data types. However, the data type TN has been deprecated in HL7 being replaced by XTN (extended telephone number). The participant's name is also referenced as an ST data type. It is unclear why the use of the current XPN data type (extended person name) is not used instead. As an example, using XPN would be extremely beneficial for the

OPMD study where many individuals have similar or identical first and last names. As a result, we record Maiden Name, which is not listed as a Standard Common Data Element. In the most up-to-date HL7 standard, the XPN data type includes an optional field for Maiden Name that would support this practice.

Although we have identified several issues with conforming to the Standard and its first revision, we applaud the efforts of the ORDR and anxiously await publication of the results from its pilot project. However, this Standard will be like all data and terminology standards in that they are never completed but rather require constant updating, expansion, and re-release. Registries that want to conform to the Standard need to take this into account when they design their work flow process. As with the OPMD registry's experience, several changes had to be made to accommodate the revisions included in going from version 1 to version 2. In addition, many projects, such as the OPMD Patient Registry, cannot wait for the "perfect" version of the standard to be released. Accordingly, we believe it would be helpful if the ORDR would provide more frequent communication or transparency into the process of the Standard as it is developing. This would be of special benefit for the projects such as the OPMD registry that are not included in their pilot project.

Despite the difficulties and limitations we outlined in this paper, there are potential benefits for implementation of the Standard across multiple patient registries. An aggregated registry of all rare disease registries will facilitate research on rare diseases that has not yet been possible. This will allow investigation of broader questions including studies of characteristics and treatments common to two or more rare diseases. For example, OPMD patients suffer from dysphagia. Individuals with myotonic dystrophy and inclusion body myositis, two other rare diseases, may also develop dysphagia. Due to lack of adequate sample sizes for statistical power, research questions related to novel dysphagia treatments cannot be easily answered by sampling from one of these patient populations alone. By recruiting participants from all three of these rare disease populations—through the use of a global rare disease registry—sufficient sample sizes could be achieved to conduct a well-powered clinical trial. Research on rare diseases may also lead to treatments for more common diseases[3]. For example, the forerunner of today's widely – used cholesterol-lowering statin medications, a drug called compactin, was first tested in individuals with the rare disease homozygous familial hypercholestrerolemia.[25] This work went on to benefit much larger groups of patients (e.g., coronary artery disease, renal failure).

Conclusion

The creation of the ORDR's registry Standard holds great potential to be a significant facilitator of research on rare diseases despite some limitations and unresolved issues. The aggregation of rare disease registries may lead to cures that benefit not only patients with a rare disease but also patients with more common diseases. No standard is perfect but adoption of the Standard by as many disease registries as possible is the crucial first step toward enabling these exciting possibilities. As Voltaire wrote in his poem La *Bégueule,* "In his writings, a wise Italian says that the best is the enemy of the good."

Reference

1. Institute of Medicine. Rare Diseases and Orphan Products: Accelerating Research and Development. Washington DC: The National Academies Press; 2010.
2. Asthma and Allergy Foundation. Asthma Overview. 2013 [cited 2013 March 1]; Available from: http://www.aafa.org/display.cfm?id=8.
3. Griggs RC, Batshaw M, Dunkle M, Gopal-Srivastava R, Kaye E, Krischer J, et al. Clinical research for rare disease: opportunities, challenges, and solutions. MGM. 2009;96(1):20-6. Epub 2008/11/18.
4. Office of Rare Disease Research. Global Rare Disease Patient Registry and Data Repository Background. 2013 [cited 2013 March 1]; Available from: http://rarediseases.info.nih.gov/Wrapper.aspx?src=asp/resources/ord_brochure.html.
5. Dreyer NAG, S. Regestries fro Robust Evidence. JAMA. 2009;302(7):790-1.
6. Gliklich RE. the emerging role of the patient registry. Good Clinical Practice Journal. 2006;December:14-7.
7. Richesson R, Shereff D, Andrews J. [RD] PRISM Library: Patient Registry Item Specifications and Metadata for Rare Diseases. J Libr Metadata. 2010;10(2-3):119-35. Epub 2010/11/09.

8. Richesson RL, Sutphen R, Shereff D, Krischer JP. The Rare Diseases Clinical Research Network Contact Registry update: features and functionality. Contemp Clin Trials. 2012;33(4):647-56. Epub 2012/03/13.
9. Sarkozy A, Bushby K, Beroud C, Lochmuller H. 157th ENMC International Workshop: patient registries for rare, inherited muscular disorders 25-27 January 2008 Naarden, The Netherlands. Neuromuscul Disord. 2008;18(12):997-1001. Epub 2008/10/25.
10. Jenders RA, McDonald CJ, Rubinstein Y, Groft SC. Applying Standards to Public Health: An Information Model for a Global Rare-Diseases Registry. AMIA 2011 Annual Symposium; Washington, D.C.2011. p. 1819.
11. Richesson RL, Nadkarni P. Data standards for clinical research data collection forms: current status and challenges. J Am Med Inf Assoc. 2011;18(3):341-6. Epub 2011/04/14.
12. Richesson RL. Data standards in diabetes patient registries. J Diabetes Sci Technol. 2011;5(3):476-85. Epub 2011/07/05.
13. Park HS, Lloyd S, Decker RH, Wilson LD, Yu JB. Overview of the Surveillance, Epidemiology, and End Results database: evolution, data variables, and quality assurance. Curr Probl Cancer. 2012;36(4):183-90. Epub 2012/04/07.
14. Rubinstein YR, Groft SC, Bartek R, Brown K, Christensen RA, Collier E, et al. Creating a global rare disease patient registry linked to a rare diseases biorepository database: Rare Disease-HUB (RD-HUB). Contemp Clin Trials. 2010;31(5):394-404. Epub 2010/07/09.
15. Office of Rare Disease Research. Global Rare Disease Patient Registry and Data Repository Common Data Elements. 2013 [cited 2013 March 1]; Available from: http://www.grdr.info/index.php?option=com_content&view=article&id=3&Itemid=5.
16. Victor M, Hayes R, Adams RD. Oculopharyngeal muscular dystrophy. A familial disease of late life characterized by dysphagia and progressive ptosis of the evelids. New Engl J Med. 1962;267:1267-72.
17. Becher MW, Morrison L, Davis LE, Maki WC, King MK, Bicknell JM, et al. Oculopharyngeal muscular dystrophy in Hispanic New Mexicans. JAMA. 2001;286(19):2437-40. Epub 2001/12/26.
18. Blumen SC, Nisipeanu P, Sadeh M, Asherov A, Tome FM, Korczyn AD. Clinical features of oculopharyngeal muscular dystrophy among Bukhara Jews. Neuromuscul Disord. 1993;3(5-6):575-7. Epub 1993/09/01.
19. Bouchard JP, Brais B, Brunet D, Gould PV, Rouleau GA. Recent studies on oculopharyngeal muscular dystrophy in Quebec. Neuromuscul Disord. 1997;7 Suppl 1:S22-9. Epub 1997/12/10.
20. Van Der Sluijs BM, Hoefsloot LH, Padberg GW, Van Der Maarel SM, Van Engelen BG. Oculopharyngeal muscular dystrophy with limb girdle weakness as major complaint. J Neurol. 2003;250(11):1307-12. Epub 2003/12/04.
21. McHorney CA, Robbins J, Lomax K, Rosenbek JC, Chignell K, Kramer AE, et al. The SWAL-QOL and SWAL-CARE outcomes tool for oropharyngeal dysphagia in adults: III. Documentation of reliability and validity. Dysphagia. 2002;17(2):97-114. Epub 2002/04/17.
22. Wallace MB, Perelman LT, Backman V, Crawford JM, Fitzmaurice M, Seiler M, et al. Endoscopic detection of dysplasia in patients with Barrett's esophagus using light-scattering spectroscopy Gastroenterology. 2000;119(3):677-82.
23. Renenstrief Institute Inc., LOINC. LOINC® users' guide. Indianapolis, IN: Regenstrief Institute; 2012.
24. ISO. Code elements for Bonaire, Saint Eustatius and Saba, Curaçao and Sint Maarten (Dutch part), update of other territories and minor corrections. 2010 [cited 2013 March 1]; Available from: http://www.iso.org/iso/iso_3166-1_newsletter_vi-8_split_of_the_dutch_antilles_final-en.pdf.
25. Endo A. A gift from nature:the birth of statins. Nat Med. 2008;14(10):xxiv-xxvi.