

# Trends of intractable disease research using biosamples

Mayako Tada<sup>1</sup>, Makoto Hirata<sup>1</sup>, Hanayuki Okura<sup>2</sup>, Satoru Yoshida<sup>2</sup> and Akifumi Matsuyama<sup>1,2</sup>

<sup>1</sup>Laboratory of Rare Disease Biospecimen, Center for Rare Disease Research, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8 Asagi, Saito, Ibaraki-city, Osaka, Japan

<sup>2</sup>Platform of Therapeutics for Rare Disease, Center for Rare Disease Research, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8 Asagi, Saito, Ibaraki-city, Osaka, Japan

## Abstract

For biosources banking one of the most important objective to support is intractable disease studies. However, it is difficult to collect enough biosources because the number of patients is small. In addition, many biobanks have a major problem that they have to store unused biosamples for a long time after collected them. The purpose of this study is to examine the trend of intractable disease research using biosamples aiming to the most efficient biobank operation with collecting and distribution of such biosamples. PubMed provides references and abstracts on life science and biological medicine which were used as search tool. The search keywords used in this study were types of biosamples (DNA, cell, tissue, e.t.c.) and 330 intractable diseases in Japan. As a result, research trends in 63 years (from Jan. 1954 through Oct. 2017) for intractable disease studies had shown that cells were the most used in research. Tissues were the second most used biosamples in research. This study is considered to support the development of efficient biobanks that have function for the collection and the distribution of incurable disease biosamples.

## Introduction

In recent years, various biobanks have been established around the world [1-3]. These biobanks have contributed significantly to advances in medical science, such as drug discovery, personalized medicine, and the elucidation of the disease mechanism.

Biobanks, although thus essential for medical development, have several problems or issues. Among others, a major issue is that many biobanks continue to store many samples that have not been distributed. This means that collected samples have not been adequately used for research. For the development of biobanks with the distribution function in place, goal-oriented remedies for banking are necessary and investigation of what types of samples are used for research on specific diseases is desirable. On the other hand, intractable diseases are defined as “(1) diseases whose cause has not yet been detected, for which there is no established therapy, and which may likely leave an aftereffect and (2) diseases which are chronic and pose not only financial problems but a heavy burden on the patients’ family due to the great deal of effort required in caring for the patients, including potential psychological burdens [4].”

Intractable disease studies have the issue that it is difficult to collect samples because of a limited number of cases and a small number of patients. Then, we have focused on intractable disease samples and research trends to help to facilitate the collection and distribution by biobanks handling intractable disease samples and information.

A questionnaire survey is one of the most commonly used survey techniques. This type of survey obtains responses to several questions, such as “What types of samples do you want to use for your intractable disease studies?” However, it may have disadvantages, including the questionable reliability of responses and a bias in respondents.

This study was designed to explore the trend of intractable disease research using biosamples from researchers around the world on whom a questionnaire survey could not have been conducted.

PubMed, which is utilized by researcher users, was used for the study and the types of samples were searched with DNA, serum, plasma, RNA, cell, and tissue as the major categories with induced pluripotent stem (iPS) [5] cells added as a medium category in the cell large category. The iPS cells have attracted a lot of attention as a therapeutic approach for Parkinson’s disease, an intractable disease [6]. We also present a discussion with a search specific to iPS cells added, although there are many intractable disease studies using cells.

## Materials and methods

A total of 330 intractable diseases designated by the Ministry of Health, Labour and Welfare [7] (enforced on April 1, 2017) were included in search. Orphanet [8], KEGG disease [9], and National Organization for Rare Disorders (NORD) [10] were consulted for English names of the designated intractable DISEASE [11]. PubMed was used as a search tool, which comprises citations from Medline, life science journals, and online books. A new program was developed to comprehensively search at least 27 million pieces of biomedical literature in PubMed for types of samples used in intractable disease studies. A part of the new program is shown in (Figure 1). Briefly, data files were downloaded from the FTP directory (<ftp://ftp.ncbi.nlm.nih.gov/pubmed/file>) using a browser. A tool was developed to summarize results from the search of titles and abstracts of all pieces of literature for the 330 diseases and types of samples.

**Correspondence to:** Director, Center for Rare Disease Research, National Institutes of Biomedical Innovation, Health and Nutrition, 7-6-8 Asagi, Saito, Ibaraki-city, Osaka, Japan, E-mail: [akifumi-matsuyama@umin.ac.jp](mailto:akifumi-matsuyama@umin.ac.jp)

**Key words:** intractable disease, rare disease, biobank, biosample, biosample distribution

**Received:** November 28, 2017; **Accepted:** December 15, 2017; **Published:** December 19, 2017

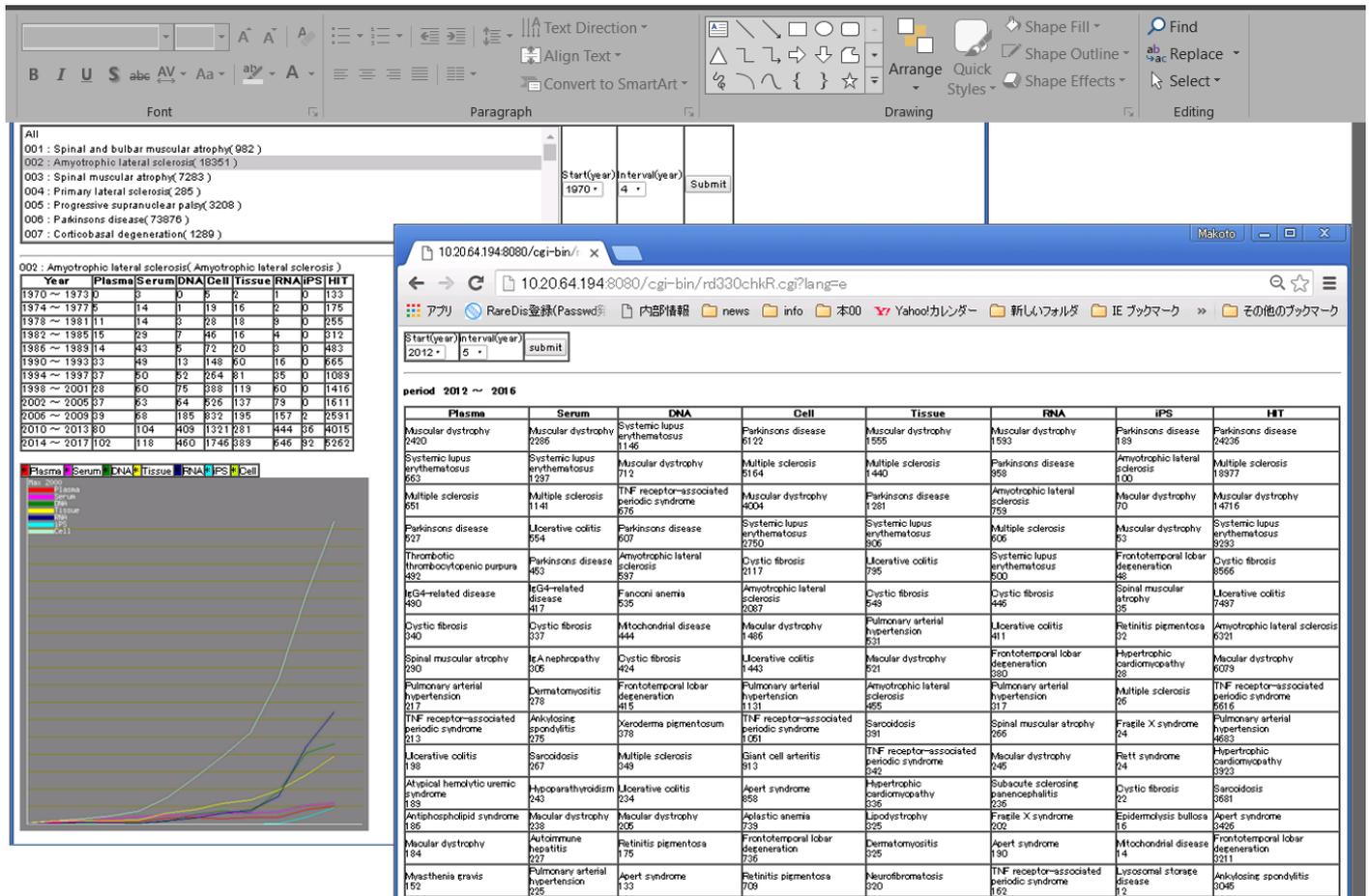


Figure 1. New Program: the search result of intractable diseases and biosamples.

Among the types of samples collected in many biobanks, the seven types of samples of DNA, serum, plasma, RNA, cell, tissue, and iPS cell were used as the search keywords. The search period was set in chronological order from January 1, 1954, to October 12, 2017. Searches involved summarization of the number of hits for the article titles and abstracts containing the keywords of “disease name” and/or “type of samples”.

First, the results of searches for each disease name of the 330 diseases and types of samples were summarized by type of sample for each year. Next, with reference to the disease groups defined by the Japan Intractable Disease Information Center, nine diseases within these groups were randomly selected from the 330 diseases (i.e., neurological/muscular system: Parkinson’s disease, immune system: systemic lupus erythematosus, visual system: retinal pigment degeneration, cardiovascular 1 disease, skin/connective tissue system: pemphigus, respiratory system: idiopathic interstitial pneumonia, endocrine system: hypoparathyroidism, bone/articular 1 disease: ankylosing spondylitis, chromosome- or gene-associated syndrome: Rubinstein-Taybi syndrome, digestive system: biliary obstruction). The types of samples and research trends were examined for the nine selected diseases. Finally, the 10 most common disease names for each type of samples were determined based on the results of search for disease name and type of sample between January 1 and October 21, 2017. Searches were conducted with the new program.

## Discussion

This study examined the trend of intractable disease research by summarizing many publications available in PubMed which had names of designated intractable diseases and types of samples in the title and abstract. Although serum was frequently used in research on the 330 intractable diseases at certain times, research using cells has recently increased markedly. In addition, iPS cells as a potential therapeutic approach for intractable diseases were found to have been increasingly used in recent years.

The trends in the use of samples observed in this study may be useful to pharmaceutical companies developing therapies for intractable diseases or intractable disease researchers. Moreover, they may be meaningful in discussing types of samples to be collected with sample provider institutions.

Collection and distributions of samples require research ethical approval [12, 13] and conclusion of material transfer agreements (MTAs). In addition, busy physicians have to spend their labor to obtain informed consent for provision of samples to biobanks from patients when they collect samples. Management of biobanks requires enormous costs to maintain freezers and liquid nitrogen tanks necessary for storage of samples. In these contexts, if a trend of research on a certain disease indicates high needs for cell or DNA samples, limiting collection to such samples may prevent plasma or serum samples less likely to be distributed in the future from coming

to nothing. Moreover, the creation of research focusing on samples with limited research outcomes is expected without getting obsessed with trends to try to conduct a highly original research.

Although this study included designated intractable diseases in Japan, 5,856 intractable diseases are now registered on Orphanet around the world (as of November 2, 2017). We intend to expand this study to include diseases registered in Orphanet. This study is considered to contribute to various approaches to intractable disease research, including development of efficient biobanks that have function for the collection and distribution of samples for patients with intractable diseases as well as for the next generation.

## Results

### Chronological changes

Changes in the types of samples used for research on the 330 designated intractable diseases showed that the number of hits for “cell” was far greater than the other types of samples from 1966 to now (Figure 2). The number of hits for “cell” included that for “iPS cells.” Serum was the next most frequently used type of sample until 1995, after which tissue replaced serum. RNA tended to be the least commonly used type of sample evaluated.

The number of hits for “iPS cell,” which has received attention in the field of intractable diseases, increased rapidly from 2008. Based on the search results in 2016, in which iPS showed the greatest number of hits, the overall number of hits generated with the 330 diseases and the keyword “cell” was 13,158 hits, including 314 hits for “iPS cell,” which accounted for 2.3% of the total number of hits for “cell.” For Parkinson’s disease in 2016, the number of hits for “cell” was 540 hits, 3.7% of which was represented by iPS cell (58 hits), suggesting that research on iPS cells in Parkinson’s disease tended to be promoted compared with research on the 330 diseases using cells.

As shown in (Figure 3), the number of hits tended to increase for any of the types of samples until 2016. Because decreases in the number of hits in 2017 were due to the search results as of October 12, 2017, being used, summarization results up to December 31, 2017, will also need to be addressed.

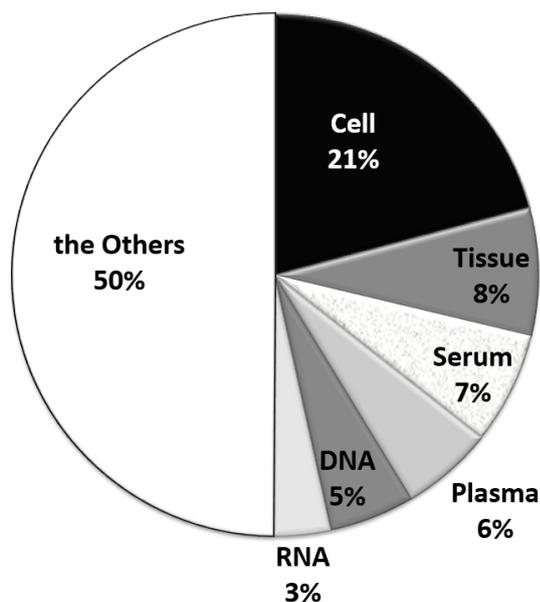


Figure 2. Percentage of research by using biosample on 330 intractable diseases research.

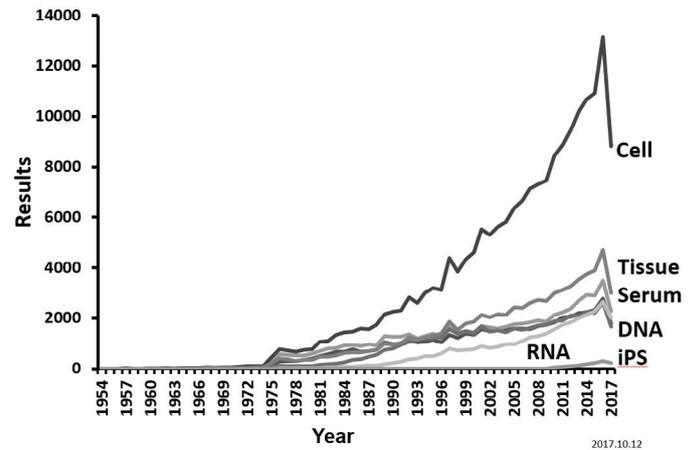


Figure 3. Percentage of each biosample in 330 intractable diseases.

### Trends of research on nine intractable diseases

The new program enabled examination of research trends until the most recent research for all the 330 diseases using varying search intervals (e.g., intervals of one year, two years, and five years) and times of the start of searches. As an example, in this study, a total of nine of the 330 diseases, each randomly selected from each of the nine disease groups of designated intractable diseases were studied to examine research trends for the types of samples (for each year between January 1, 1954, and October 12, 2017). Although the total number of hits for “cell” across the 330 diseases shown in (Figure 3) was predominantly greater, the number of hits for “cell” was greater in five of the nine randomly selected diseases, and the types of samples did not show markedly high numbers of hits in the other diseases (Figure 4).

The graphs show changes from the start year after 1954 onward, in other words, the year for which at least one hit was generated even though it was the hit for disease names alone. As shown in (Figure 3), the diseases studied since the 1950s included “6: Parkinson’s disease,” “35: pemphigus,” and “49: systemic lupus erythematosus” (numbers indicate the notice numbers for designated intractable diseases). “85: idiopathic interstitial pneumonia” was found to have been more recently started to be studied (since the 1970s). Thus, the background to the research could be understood for the diseases, including “26: HTLV-1 associated myelopathy” in the neurological system group (not shown in the figure), which has been studied since 1987 and is estimated to affect 3,000 individuals in Japan. In addition, “102: Rubinstein-Taybi syndrome,” one of the nine diseases, showed a very small number of hits of not more than 5, suggesting that reported studies were limited by the rare nature of this disease. Thus, marked differences in the number of hits were found among the diseases.

### Most common intractable diseases by type of samples

Largest numbers (numbers under disease names indicate the number of hits) of hits for “disease name” and “type of samples” from January 1 to October 21, 2017 (it should be noted that the dates of this search were different from other dates of the search because data had been updated from October 12) (Table 1). For the six types of samples (with iPS cell included in cell), 45% of the 10 most common diseases represented neurological/muscular diseases. This result was not reflective of overall ratio with 83 of the 330 diseases (25%) being neurological/muscular diseases, suggesting that these six types of samples played an important role in research on intractable neurological/muscular diseases.

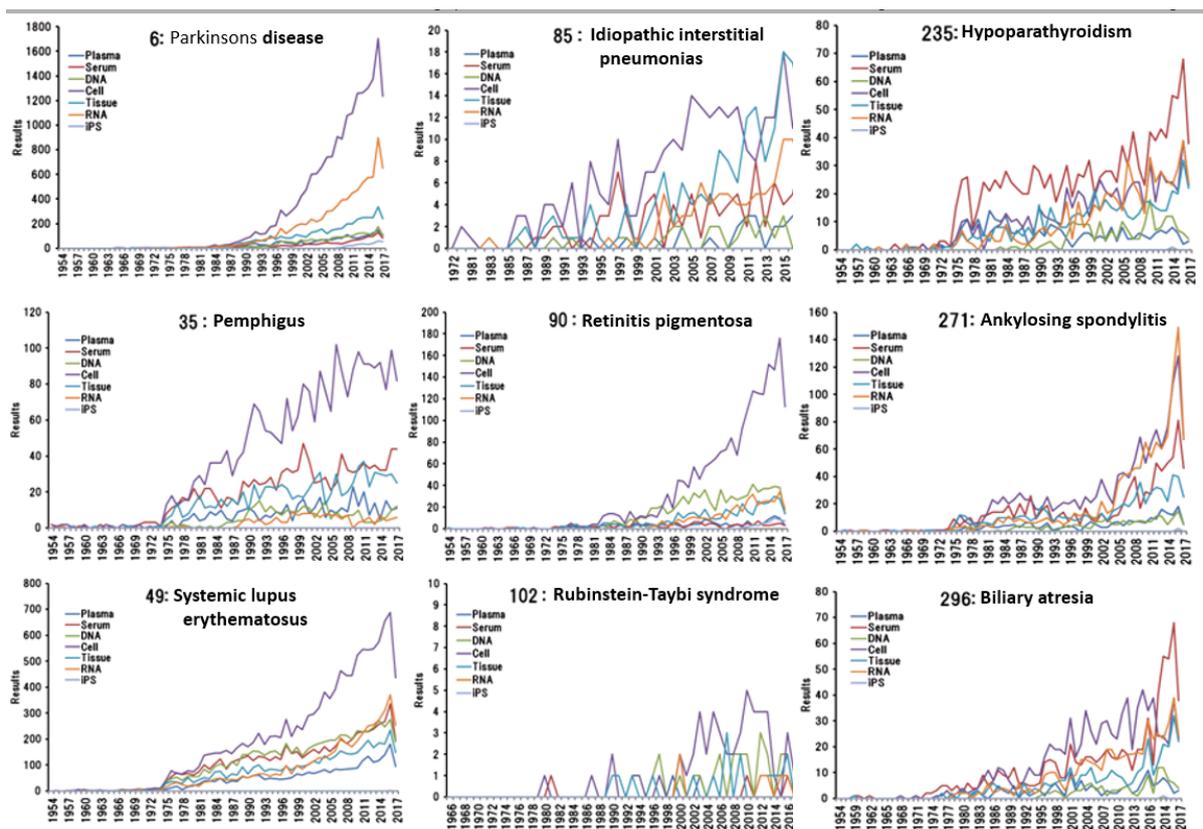


Figure 4. Trend of randomly extracted 9 diseases from 330 intractable diseases by biosamples.

Table 1. Top10 of intractable disease research by using biosamples (The red: neuromuscular system, The black: The others system, Upper: intractable disease, Lower: search results).

Ranking	Plasma	Serum	DNA	Cell	Tissue	RNA	iPS	Only disease (No biosample)
1	Muscular dystrophy 350	Muscular dystrophy 375	Systemic lupus erythematosus 176	Parkinsons disease 1122	Parkinsons disease 243	Muscular dystrophy 287	Parkinsons disease 52	Parkinsons disease 4517
2	Multiple sclerosis/ Optic neuromyelitis 112	Multiple sclerosis/ Optic neuromyelitis 210	TNF receptor-associated periodic syndrome 140	Multiple sclerosis/ Optic neuromyelitis 894	Multiple sclerosis/ Optic neuromyelitis 234	Parkinsons disease 217	Amyotrophic lateral sclerosis 20	Multiple sclerosis/ Optic neuromyelitis 3239
3	Parkinsons disease 111	Systemic lupus erythematosus 198	Muscular dystrophy 112	Muscular dystrophy 645	Muscular dystrophy 231	Amyotrophic lateral sclerosis 151	Macular dystrophy 18	Muscular dystrophy 2364
4	Systemic lupus erythematosus 99	Ulcerative colitis 95	Amyotrophic lateral sclerosis 102	Systemic lupus erythematosus 416	Systemic lupus erythematosus 149	Multiple sclerosis/ Optic neuromyelitis 132	Muscular dystrophy 15	Ulcerative colitis 1552
5	IgG4-related disease 73	Parkinsons disease 82	Parkinsons disease 99	Amyotrophic lateral sclerosis 390	Ulcerative colitis 137	Ulcerative colitis 95	Frontotemporal lobar degeneration 14	Systemic lupus erythematosus 1487
6	Thrombotic thrombocytopenic purpura 70	IgG4-related disease 63	Fanconi anemia 78	Macular dystrophy 309	Macular dystrophy 90	Systemic lupus erythematosus 83	Hepatic glycogenesis 7	Cystic fibrosis 1333
7	TNF receptor-associated periodic syndrome 44	IgA nephropathy 53	Multiple sclerosis/ Optic neuromyelitis 70	Cystic fibrosis 302	Amyotrophic lateral sclerosis 87	Cystic fibrosis 72	Retinitis pigmentosa 7	Amyotrophic lateral sclerosis 1163
8	Spinal muscular atrophy 44	Cystic fibrosis 48	Mitochondrial disease 69	Ulcerative colitis 248	Cystic fibrosis 83	Spinal muscular atrophy 68	Hypertrophic cardiomyopathy 7	TNF receptor-associated periodic syndrome 1145
9	Pulmonary arterial hypertension 43	Ankylosing spondylitis 46	Frontotemporal lobar degeneration 59	TNF receptor-associated periodic syndrome 229	TNF receptor-associated periodic syndrome 80	Frontotemporal lobar degeneration 67	Cystic fibrosis 6	Macular dystrophy 1067
10	Cystic fibrosis 41	Pemphigus 44	Xeroderma pigmentosum 49	Pulmonary arterial hypertension 195	Pulmonary arterial hypertension 79	Macular dystrophy 58	Rett syndrome 5	Pulmonary arterial hypertension 815

## Conclusion

These six types of samples played a vital role in research on intractable neurological/muscular diseases.

## References

1. UK Biobank (2012) Protocol for large-scale prospective epidemiological resources <http://www.ukbiobank.ac.uk/wp-content/uploads/2011/11/UK-Biobank-Protocol.pdf>
2. Kuriyama S, Yaegashi N, Nagami F, Arai T, Kawaguchi Y, et al. (2016) The Tohoku Medical Megabank Project: Design and Mission. *J Epidemiol.* 26: 493-511
3. Cho SY, Hong EJ, Nam JM, Han B, Chu C, et al. (2012) Opening of the national biobank of Korea as the infrastructure of future biomedical science in Korea. *Osong Public Health Res Perspect* 3: 177-184.
4. Principles of policy for intractable diseases published by the Ministry of Health and Welfare (at that time) in 1972.
5. Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126: 663-676.
6. Kikuchi T, Morizane A, Doi D, Magotani H, Onoe, et al. (2017) Human iPS cell-derived dopaminergic neurons function in a primate Parkinson's disease model. *Nature.* 548: 592-596, 2017
7. MHLW website. Designated intractable diseases implemented on April 1, 2017 (new/updated). <http://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000085261.html>
8. The portal for rare diseases and orphan drugs. Website. <http://www.orpha.net/consor/cgi-bin/index.php?lng=EN>
9. KEGG DISEASE. Website. [http://www.genome.jp/kegg/disease/disease\\_ja.html](http://www.genome.jp/kegg/disease/disease_ja.html)
10. NORD National Organization for Rare Disorders. Website. <https://rarediseases.org/>
11. Suzuki M, Sakate R, Fukagawa A (2016) Office of Pharmaceutical Industry Research (OPIR) News. *OPIR view and actions* 49: 35-46.
12. Tsutsumi M (2014) Outline of revised ethical guidelines for human genome/gene analysis research in Japan. *Organ Biology* 21: 9-15.
13. Guidance of Ethical Guidelines for Medical and Health Research Involving Human Subjects. February 9, 2015 (partially revised on May 29, 2017)