Do geneticists need Babel fish?

The Babel fish is small, yellow and leech-like and is probably the oddest thing in the Universe. It feeds on brainwave energy received not from its own carrier but from those around it. It absorbs all unconscious mental frequencies from this brainwave energy to nourish itself, then excretes into the mind of its carrier a telepathic matrix formed by the signals from the speech centers of the brain which supplied them.

The practical upshot of all this is that if you stick a Babel fish in your ear you can instantly understand anything said to you in any sort of language.[1]

Geneticists need to understand each other, and sometimes it isn't easy. The recent remarkable increase in knowledge about our genome, genetic mutations that may or may not be associated with disease, and genetic diseases that may or may not be associated with particular mutations has made the language we have conventionally used in genetics clinics and laboratories increasingly inappropriate and confusing. Robin and Biesecker are to be congratulated for bringing this problem to our attention and starting us on the way to a solution.[2,3]

As clinicians, we are aware that "labeling" patients can be both good and bad. On the bad side, labeling can sometimes produce diminished expectations or expectations of failure that may be unfair to an individual who has been diagnosed with a condition associated with serious intellectual or behavioral deficiencies. Labeling can also impair medical care if symptoms of an unrelated condition are ignored as "just part of the syndrome" instead of being investigated as thoroughly as they might be under other circumstances.

There can also be important advantages to proper diagnostic labeling, and this is what the proposal by Robin and Biesecker is all about. The intent of proper diagnostic labeling is to facilitate more effective communication among affected individuals, their families, and those who provide them with medical care, education, or social services. Proper diagnostic labeling can help establish more appropriate expectations, improve preventive health care, assist in the design of optimal medical and educational interventions, and aid in obtaining necessary services. In addition, proper diagnostic labeling is prerequisite to accurate genetic counseling—the most common serious error made in genetic counseling is providing the right information for the wrong diagnosis.

Proper diagnostic labeling is also essential for research. The power of genotype–phenotype correlations to provide new insights into disease pathogenesis or gene function is absolutely dependent on proper diagnostic labeling. Natural history studies, trials of new therapeutic modalities, genetic epidemiology, and recognition of gene–environment interactions all require accurate clinical diagnosis and proper diagnostic labeling.

Robin and Biesecker[3] provide one suggestion about how this labeling might be done. Others have made different proposals about how to join clinical and molecular genetic information into a single diagnostic labeling system.[4] I think that an effective system of nomenclature for genetic phenotypes (including those that are not "purely" genetic) needs to include more than just a systematic and flexible approach to choosing names. Several other comprehensive systems of genetic or clinical nomenclature have been developed over the past several years,[5–9] and we should learn from the experience gained in these other contexts in developing a nomenclature for clinical genetics. Therefore, I believe that in addition to the desirable characteristics of a clinical genetics nomenclature system listed by Robin and Biesecker, such a system should:

- Not attempt to incorporate *all* information about a patient, but only the information that is *essential* to establish a phenotypic diagnosis as precisely as possible.
- Produce consistent designations for identical patients seen by different clinicians or in different centers.
- Be amenable to convenient computerization and integration with other relevant computerized databases.

Inclusion of these criteria requires substantial modification of the proposal put forward by Robin and Biesecker and provides important insights into how the system should be applied in practice.

According to Robin and Biesecker, the purpose of this system is "to accurately describe a phenotype in a particular patient." It is obvious that certain information is irrelevant to this purpose (e.g., the patient's name) and that certain other information is essential (e.g., the presence of a typical dysmorphic syndrome). Robin and Biesecker argue, correctly I think, that etiological information is sometimes useful in characterizing the phenotype, but I would draw the circle of relevance more tightly than they do. I do not believe that being at increased risk for having a child with a chromosomal abnormality because of the results of a serum triple screen is a diagnostic phenotype. The results of a triple screening test do not describe the underlying molecular genetic defect or environmental or nongenetic factors related to any abnormal phenotype and, therefore, do not belong in this nomenclature. Similarly, I do not think that negative molecular genetic studies are generally useful for characterizing a phenotype; a "negative study" means different things in different labs and even in the same lab at different times. I hasten to add that I believe it is always useful to keep track of this information about a patient; I just do not think that it should be considered part of the phenotypic diagnosis in most instances. There are a few cases in which a typical phenotype is found without the mutation that is almost always responsible for it, e.g., achondroplasia without a Gly380Arg

amino acid substitution in fibroblast growth factor receptor 3 or fragile X syndrome without expansion of the CGG trinucleotide repeat in *FMR1.* These cases should be treated as exceptions. The *presence* of a particular pathogenic mutation, on the other hand, is probably always worth including in the diagnostic designation because this information may be useful for genotype–phenotype correlations in the future even if such associations are not apparent now.

Inclusion of environmental etiology codes is likely to be problematic in many instances as well. It is one thing to recognize in an epidemiological study that a certain environmental exposure during pregnancy (e.g., cigarette smoking) is *associated* with a particular congenital anomaly (e.g., cleft lip and palate). It is quite another thing to say that this exposure had something to do with the development of the congenital anomaly in an individual child. The same difficulty exists in trying to attribute phenotypic manifestations in an individual patient to a genetic polymorphism that may be associated with a 2-fold increased risk in a population study. It seems doubtful that the clinical care of a patient or family would be affected by such an attribution, so I would oppose including this information in a phenotypic nomenclature system. Again, there are uncommon exceptions, but these should be treated in an exceptional manner. For example, the environmental etiology of fetal alcohol syndrome affects the care of an individual with this condition in many ways and should be part of the nomenclature.

Consistency of diagnostic coding is an essential feature of any useful system of nomenclature. If similar cases are not coded in a similar way by different coders, the system is unlikely to improve communication among professionals and cannot be used effectively to identify patients for research. This is a clear message from the experience with diagnostic coding in other genetic or clinical contexts and explains why all of these systems include extensive and precise rules about how to code anomalies.[5–9] These other systems also use a limited set of terminology to name abnormalities. This is sorely needed in medical genetics. It would make things much easier for patients, parents, health and other professionals, and researchers if the phenotypic manifestations of multiple hereditary exostoses, for example, were always called that and not diaphyseal aclasis, multiple osteocartilaginous exostoses, osteochondromatosis, multiple cartilaginous exostoses, or multiple osteochondromatosis. A system of standardized clinical genetic nomenclature that is flexible enough to accommodate exceptions, uncertainty, and change would be very valuable in clinical genetics.

A related issue is the requirement for clinical genetics nomenclature to be amenable to convenient computerization and integration with other relevant systems and databases. Small issues of design, such as the use of particular punctuation marks to separate fields or entries within a field, can make a huge difference in a computer-based system's ability to retrieve data efficiently. Computerizing the system from the beginning will facilitate accurate data entry and permit the arcane numerical or alphanumeric codes that are usually employed to be concealed within the software. This would enable users to work in any language with which they are familiar and allow standard diagnostic designations to be generated, stored, and printed out automatically. The software could incorporate context-specific prompts (e.g., asking about certain kinds of molecular genetic testing when a particular diagnosis is entered) and provide useful links to various clinical, administrative, laboratory, and research databases. These links would allow all of the other information we need to know about a patient—negative laboratory tests, history of exposures during gestation, information about screening tests that modify risks, and family history, among others—to be available instantly in association with the diagnostic designation without actually having to be part of it.

Robin and Biesecker put their proposal forward to stimulate discussion of a standard system of phenotypic nomenclature for clinical genetics. They have highlighted an important problem that is likely to get worse as our knowledge of the molecular genetics of both rare and common diseases improves. The American College of Medical Genetics should accept Robin and Biesecker's challenge and convene an international expert committee to develop an appropriate system of clinical genetics nomenclature. Failure to do so now will only increase the likelihood that we shall have to turn to Babel fish to help us understand each other in the future.

J.M. Friedman, MD, PhD
Department of Medical Genetics
University of British Columbia
Vancouver, British Columbia
Canada

References

1. Adams D. The hitch hiker's guide to the galaxy: a trilogy in five parts. London: W. Heinemann, 1995.
2. Biesecker LG. Lumping and splitting: molecular biology in the genetics clinic. *Clin Genet* 1998;53:3–7.
3. Robin NH, Biesecker LG. Considerations for a multiaxis nomenclature system for medical genetics. *Genet Med* 2001;3:290–293.
4. Cohen MM, MacLean RE. Should syndromes be defined phenotypically or molecularly? Resolution of the dilemma. *Am J Med Genet* 1999;86:203–204.
5. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 4th ed, text revision. Washington, DC: American Psychiatric Association, 2000.
6. Antonarakis SE. Recommendations for a nomenclature system for human gene mutations: Nomenclature Working Group. *Hum Mutat* 1998;11:1–3.
7. Ct RA, College of American Pathologists Committee on Nomenclature and Classification of Disease. Systematized nomenclature of medicine. Skokie, IL: College of American Pathologists, 1976.
8. Mitelman F. ISCN: an international system for human cytogenetic nomenclature. Basel: S. Karger, 1995.
9. World Health Organization. ICD-10: International statistical classification of diseases and related health problems, 10th revision. Geneva: World Health Organization, 1992–1994.