What's in a Name? How Is Nomenclature Developed and How Can I Keep Up?

Chair:
Cheryl Iverson
American Medical Association
Chicago, Illinois

Panelists:
Stuart Nelson
National Library of Medicine
Bethesda, Maryland

John M Opitz
American Journal of Medical Genetics
Salt Lake City, Utah

Reporter:
Deborah Flapan
JAMA Medical News
Chicago, Illinois

Keeping up with the rapid name changes in medicine can be a full-time job. Stuart Nelson and John M Opitz both expressed frustration with current naming systems.

Nelson, head of medical subject headings at the National Library of Medicine (NLM), described how he determines medical subject headings (MeSH), the "controlled vocabulary" through which NLM catalogs and indexes medical literature. To be a subject heading for MeSH, said Nelson, a name must be current, accurate, and usable and must partition the information space. Names should be expressive, having a distinct meaning that does not depend on context, and should be unique, such as the "reaper gene", which appears during apoptosis (cell death). But some creators of names go too far and want their names to be flashy or interesting, accuracy no longer

being their main concern. Ultimately, usage determines which names are adopted; editors have a large role to play in that regard.

Nelson's particular concern is the method of naming genes. A 3-letter combination is used, and whether the gene name is italicized or has an initial capital letter can change which gene is in question. Nelson pointed out that the human genome comprises at least 100 000 genes, but the 3-letter system of naming genes allows for only 4732 combinations. With so few combinations possible, uniqueness of names will be impossible to maintain. Furthermore, with the current convention of italicizing gene names but not gene products, it is difficult in many computer systems to determine whether a gene or the gene product is being discussed.

Nelson's wish list for editors to help maintain nomenclature included the absence of unidentified acronyms; the use of expressive names, not ones that depend on context to determine meaning; and the avoidance of obscure, inaccessible jargon.

In genomics, the best way to stay abreast of nomenclature, Nelson said, is to use the Internet.

Opitz said that he is also aware of sweeping changes in nomenclature because of the various roles he plays, including being the editor of the American Journal of Medical Genetics. The 4 most common conventions for naming in the medical literature, he said, are eponyms (which are best used in the nonpossessive form because most creators of names do not have the syndrome they're naming and do not own the name itself), acronyms, patients' surnames, and numbers.

Keeping track of nomenclature developments is especially difficult when new terms are added to the literature daily, Opitz said, echoing a statement by Nelson.

Opitz also concurred with Nelson regarding the necessity of a new and consistent method for naming genes. Opitz said that the recent propensity for "snappy" names should be discouraged because these names are often used for effect rather than for effectiveness and understanding; such names are often not specific enough and therefore are not helpful, and sometimes they are offensive, such as the "arse syndrome".

Opitz's recommendations for nomenclature include creation of one system of symbols for all genes (for example, both human and mouse). He would purge journal articles of all technical jargon except terms defined in an appendix, and he would publish only articles "written in clear prose, understandable by a reasonably well-educated layperson." Doing that, he said, would guarantee that editors would get articles of the desired quality, and it would solve complexity problems. It would also address the issue of the growing number of patients and their families who have become more active in the diagnostic process because of easy access to medical information via the Internet.

Internet resources for tracking nomenclature (last accessed 17 May 1998):

Genome databases:

National Center for Biotechnology Information:
www.ncbi.nlm.nih.gov/

Saccharomyces Genome Database:
genome-www.stanford.edu/saccharomyces/

Flybase: A Database of the Drosophila Genome:
flybase.bio.indiana.edu:82/

The Galton Laboratory (human genetics and evolutionary biology):
www.gene.ucl.ac.uk/

The Genome Database (Human Genome Project):
gdbwww.gdb.org/

The Jackson Laboratory: Mouse Genome Informatics:
www.informatics.jax.org/

Other resources:

SWISS-PROT: Annotated Protein Sequence Database:
www.expasy.ch/sprot/sprot-top.html/

DbBrowser: The University College London's Bioinformatics Server:
www.biochem.ucl.ac.uk/bsm/dbbrowser/

The Protein Data Bank:
www.pdb.bnl.gov/

Structural Classification of Proteins (release 1.37):
scop.mrc-lmb.cam.ac.uk/scop/