

NCI Thesaurus property definitions as of September 15, 2009

Property	Use/Meaning
Accepted_Therapeutic_Use_For	Indicates a disease or condition for which this drug is an accepted treatment. Used in Chemicals and Drug Kind.
ALT_DEFINITION	English language definitions of what a source other than NCI means by the concept. These are limited to 1024 characters. They include information about the definition's source in a form that can easily be interpreted by software.
ALT_LONG_DEFINITION	Used to house the full text of a non-NCI definition that is too long to conform to the 1024 character limit of an ALT_DEFINITION.
BioCarta_ID	Indicates the BioCarta ID code for a pathway. It can be used to link a BioCarta pathway reference to the primary pathway information.
CAS_Registry	Indicates the Chemical Abstracts Service number for a Chemical or Drug.
Chemical_Formula	Contains the concise expression about the atoms that constitute a particular chemical compound.
Concept_Status	Used to indicate the standing of a concept in relation to currently accepted classifications and concepts. In NCI Thesaurus concept status subtype indicates concepts with unusual and problematic characteristics that should be evaluated by people and/or programs before those concepts are used.
Contributing_Source	Used to indicate when a non-EVS entity has contributed to, and has a stake in, a concept. This is used where such entities, within or outside NCI, have indicated the need to be able to track their own concepts. A single concept can have multiple instances of this property if multiple entities have such a defined stake.
Def_Curator	Indicates that the named special project or source has primary responsibility for the creation and maintenance of the NCI definition for this concept.
DEFINITION	English language definitions of what NCI means by the concept. These are limited to 1024 characters. They may also include information about the definition's source and attribution in a form that can easily be interpreted by software.
DesignNote	Notations made by NCI vocabulary editors. They are intended to provide supplemental, unstructured information to the user or additional insight about the concept.
Display_Name	Provides an alternative Preferred Name for use in some NCI systems.
Editor_Note	Provides a place for an editor to record information that another editor may need to know about how a concept was modeled. These are used for internal purposes and are not published.
EntrezGene_ID	Indicates the accession number for a Protein or Gene within the Entrez Gene database.

Essential_Amino_Acid	Indicates that an amino acid is essential and must be included in the diet.
Essential_Fatty_Acid	Indicates a fatty acid that is not efficiently synthesized by humans and must be included in the diet.
Extensible_List	Attribute that defines whether new concepts can be added to a list or subset of terms.
FDA_Table	Flags terms that are part of an FDA data standard manual, including Route of Administration, Dosage Form, Package Type and Potency.
FDA_UNII_Code	Contains the FDA Unique Ingredient Identifier, a key component of the new federal drug information model.
FULL_SYN	Fully qualified synonym, containing the string, term type, source, and an optional source code if appropriate. Each subfield is delimited to facilitate interpretation by software.
GenBank_Accession_Number	Indicates the accession number for a Gene within the GenBank database.
Gene_Encodes_Product	Indicates the protein or RNA product expressed by a gene.
GO_Annotation	Contains information that can be used for cross-referencing to relevant entries in the Gene Ontology database including term name, code, and evidence code.
Homologous_Gene	Indicates genes from other organisms that are similar in sequence and provide insight to the gene's function.
ICD-O-3_Code	Indicates the matching ICD-O-3 code for the NCI Thesaurus concept.
Image_Link	Provides a link to an image in an external database.
In_Clinical_Trial_For	Indicates a disease for which this drug is currently being tested as a potential treatment.
IND_Code	Investigational New Drug Code assigned by the FDA to a new drug or device before entering a clinical trial; it is utilized in some NCI systems.
INFOODS	Uppercase abbreviations assigned by the International Network of Food Data Systems.
KEGG_ID	Indicates the KEGG ID code for a pathway. It can be used to link a KEGG pathway reference to the primary pathway information.
Locus_ID	Indicates the accession number for a Protein or Gene within the LocusLink database.
LONG_DEFINITION	Houses the full text of an NCI definition that is too long to conform to the 1024 character limit of a DEFINITION.
Macronutrient	Indicates a class of food components that contribute energy to the diet and that are typically consumed in multi-gram amounts.

Maps_To_LASH	Maps from the IMT disease codes used by NCI's Cancer Therapy Evaluation Program (CTEP) to the LASH morphology and topography. This is not an exact mapping.
Merge_Into	Used when two concepts are merged. The losing concept goes into the retired kind and this property is created in the losing concept to record which concept won the merge.
Merge_Source	Used temporarily in a concept that has been pre-merged, but before it has been approved for merge by a manager. This property is created in the winning concept to indicate which concept is going to be merged to it. When the actual merge occurs, this property is deleted from the concept.
Merge_Target	Used temporarily in a concept that has been pre-merged, but before it has been approved for merge by a manager. The Merge_Target is created in the losing concept to indicate which concept it needs to be merged with. When the actual merge occurs, this property is deleted from the concept.
MGI_Accession_ID	Identifier used by Jackson Laboratories Mouse Genome Informatics group to label inbred mouse strains maintained by Jackson Laboratory and other things such as mouse genes and alleles.
Micronutrient	Indicates a class of food components required for proper health and that are typically consumed in microgram or milligram amounts.
miRBase_ID	Permits attribution and external referencing of concepts cross-referenced in miRBase, allowing users to obtain provenance information from the miRBase web site.
Mitelman_Code	Felix Mitelman's codes for the topography and morphology terms that he uses to code his Chromosome Aberrations Database; data which is included in NCI's Cancer Genome Anatomy Project (CGAP) databases.
NCBI_Taxon_ID	Establishes correspondence between the NCI Thesaurus concept representing a taxon (species) and the NCBI/NLM Species Taxonomy ID assigned by the NCBI to that taxon.
NCI_META_CUI	Contains a Concept Unique Identifier for those concepts that appear in NCI Metathesaurus but not in the NLM UMLS.
NDFRT_Code	Holds the unique US Veterans Affairs Administration National Drug Formulary Reference Terminology concept code associated with the NDFRT_Name.
NDFRT_Name	Holds US Veterans Affairs Administration National Drug Formulary Reference Terminology concept name with the same intended meaning as the associated Thesaurus Concept. Used for the purpose of mapping between NDFRT and NCI Thesaurus.

Neoplastic_Status	Contains information about the cancer status of individual cancer concepts: Malignant, Benign, Precancerous, Unknown Malignancy Potential and Undetermined.
NSC_Code	A numeric identifier for substances submitted to the National Cancer Institute (NCI) for testing and evaluation. It is named after the Cancer Chemotherapy National Service Center (CCNSC), created in 1955, which originally ran the testing program and assigned these numbers. The testing program and repository is now run by the NCI Developmental Therapeutics Program (DTP). Assignment of NSC numbers was never done in a way that would guarantee one and only one NSC number would be assigned to a particular substance.
Nutrient	Indicates that a food component must be taken into the body preformed and in sufficient quantity to meet the body's needs for growth, reproduction and/or maintenance of health.
OID	Stores the Object Identifier of a concept.
OLD_CHILD	Stores the name of children concepts present at the time a concept was retired.
OLD_KIND	Stores the name of the concept's kind at the time a concept was retired.
OLD_PARENT	Stores the name of parent concepts present at the time a concept was retired.
OLD_ROLE	Stores the name and value of asserted roles present at the time a concept was retired.
OLD_STATE	Stores the state of a concept, primitive or defined, at the time a concept was retired.
OMIM_Number	Indicates the accession number for a Protein or Gene within the OMIM database.
PDQ_Closed_Trial_Search_ID	Identifier used by NCI's Cancer.gov web site to retrieve PDQ Clinical Trials data for closed trials that involved specific clinical trials agents. While the identifier is the same for both open and closed trials, two properties are used to support requirements of the current browser interface, as well as the possibility of providing values only where trials will be retrieved.
PDQ_Open_Trial_Search_ID	Identifier used by NCI's Cancer.gov Web site to retrieve PDQ Clinical Trials data for open trials involving specific clinical trials agents. While the identifier is the same for both open and closed trials, two properties are used to support requirements of the current browser interface, as well as the possibility of providing values only where trials will be retrieved.
Preferred_Name	Word or phrase that NCI uses by preference to refer to the concept.
PubMedID_Primary_Reference	Used by the MMHCC vocabulary to provide a primary reference for descriptions of mouse diseases.

Related_Lash_Concept	Direct mapping from LASH (ELC2001) to the NCI Thesaurus. This property is provided for use by the Cancer Genome Anatomy Project (CGAP) and the Cancer Molecular Analysis Project (CMAP).
Related_MedDRA_Code	Nearest matching MedDRA code for MedDRA/ CTEP disease names used in coding CTEP Clinical Trials.
Relative_Enzyme_Activity	Indicates drug efficacy of the allelic product.
Semantic_Type	Describes the sort of category to which a concept belongs in the context of the UMLS semantic network.
SNP_ID	Indicates the dbSNP id number for the genetic polymorphism due to base substitutions.
Split_From	Indicates when a concept is split into two concepts. The newly created concept is given this property to record which concept it used to be part of.
Subsource	Used by MMHCC to identify the source of various mouse disease concepts.
Swiss_Prot	Indicates the accession number for a Protein within the Swiss-Prot database.
Synonym	Ways of referring to a concept that are valid alternatives to the preferred name that NCI uses to refer to the concept.
Tolerable_Level	Indicates that the Food and Nutrition Board of the National Academy of Science has established a highest level of daily intake for this food component that is likely to pose no risk of adverse health effects for almost all individuals in the general U.S. population.
UMLS_CUI	Concept unique identifiers assigned by the National Library of Medicine (NLM). If a concept in any NCI-maintained knowledgebase exists in the NLM Unified Medical Language System (UMLS), NCI includes the NLM CUI among the information we provide about the concept.
Unit	Unit of measure in which a nutrient or food component is typically expressed, ex., gram (g), minigram (mg), microgram (mcg), joule, or kilojoule.
US_Recommended_Intake	Indicates that the Food and Nutrition Board of the National Academy of Science has established an average daily intake for this nutrient that is sufficient to meet the nutritional requirements of most individuals in the general U.S. population.
USDA_ID	Numeric ID assignment by United States Department of Agriculture.
Use_For	Indicates a term that had been used in a coding system and then subsumed by the given NCI concept.