

# NCI Thesaurus property definitions as of April 4, 2006

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 from non-NCI sources describing the meaning of the concept. These are limited to 1024 characters. They also include information about the definition's source that can easily be interpreted by software.
ALT_LONG_DEFINITION	Used to house the full text of a definition from a non-NCI source that is too long to conform to the 1024 character limit of an ALT_DEFINITION. They also include information about the definition's source that can easily be interpreted by software.
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.
Concept_Status	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 concept are used.
Contributing_Source	This property is 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.
DEFINITION	English language definitions of what NCI means by the concept. These are limited to 1024 characters. They also include information about the definition's source that can easily be interpreted by software.
DesignNote	Design notes are notations made by NCI vocabulary curators. 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.
FDA_Table	Used to flag 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, contains 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 a value that can be used for cross-referencing to relevant entries in the Gene Ontology database.
Homologous_Gene	This indicates genes from other organisms that are similar in sequence and provide insight to the gene's function.
Image_Link	Provides a link to an image in an external database.

IND_Code	Investigational New Drug Code. Code assigned by the FDA to a new drug or device before entering a clinical trial; it is utilized in some NCI systems.
In_Clinical_Trial_For	Indicates a disease for which this drug is currently being tested as a potential treatment.
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.
LONG_DEFINITION	Used to house the full text of a definition that is too long to conform to the 1024 character limit of a DEFINITION.
Locus_ID	Indicates the accession number for a Protein or Gene within the LocusLink database.
Maps_To_LASH	Used for mapping 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.
MGI_Accession_ID	This property contains the identifiers used in Mouse Genome Informatics (MGI) databases to label data on gene characterization, nomenclature, mapping, gene homologies among mammals, sequence links, phenotypes, allelic variants and mutants, and strains nomenclature.
Mitelman_Code	Felix Mitelman's codes for the topography and morphology terms he uses to code his Chromosome Aberrations Database, data which is included in NCI's Cancer Genome Anatomy Project (CGAP) databases.
NCBI_Taxon_ID	This property is used to establish correspondence between the NCI Thesaurus concept representing a taxon (species) and the NCBI/NLM Species Taxonomy ID assigned by the NCBI to that taxon. The property value is a string value corresponding to integer obtained from NCBI.
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	Used to hold 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.
NSC_Code	Drug codes provided by NCI's Developmental Therapeutics Program (DTP).
OLD_CHILD	Used to store the name of children concepts present at the time a concept was retired.
OLD_KIND	Used to store the name of the concept's kind at the time a concept was retired.
OLD_PARENT	Used to store the name of parent concepts present at the time a concept was retired.
OLD_ROLE	Used to store the name and value of asserted roles present at the time a concept was retired.
OLD_STATE	Used to store 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	This property contains an identifier used by NCI's Cancer.gov Web site to retrieve PDQ Clinical Trials data for specific clinical trials agents. This property is used to retrieve closed clinical trials.
PDQ_Open_Trial_Search_ID	This property contains an identifier used by NCI's Cancer.gov Web site to retrieve PDQ Clinical Trials data for specific clinical trials agents. This property is used to retrieve open clinical trials.

Preferred_Name	The preferred name is the 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	The semantic type describes the sort of thing or 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.
Subsource	Currently 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	Synonyms are ways of referring to a concept that are valid alternatives to the preferred name that NCI uses to refer to the concept.
UMLS_CUI	Concept Unique Identifiers, or CUIs, are concept numbers 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.