

Generating NCI Thesaurus Concept Lineage

Sept 21, 2022

Supporting Artifact for Project “Expanding UMLS SemNet for COVID-19 CDEs”
(20X122QT3 Task-Order)


Wendy Ver Hoef, Samvit Solutions

Overview

- As part of the overall project, COVID CDEs have been mapped to UMLS concepts and NCIIt concepts
- To see the hierarchical relationship of the NCIIt concepts and identify common parents/ancestors, it is necessary to obtain the full concept lineage.
 - This lineage is derivable from the NCIIt concept data because it includes the child and parent concepts for all 175K+ concepts in NCIIt.
 - By consecutively looking up the parent concept, grandparent, great grandparent, etc., until we reach the root, we can automate the compilation of the lineage.
 - The same process could be automated via a variety of coding means, but Excel required little testing and only commonly available skills.
 - The purpose of this slide deck is to outline the steps used in Excel so it can be replicated logically in code or updated for future versions in the same way with Excel for the time being.

Step 1: From the NCI Thesaurus Downloads page, download the tab-delimited format

<https://evs.nci.nih.gov/evs-download/thesaurus-downloads>

**NATIONAL CANCER INSTITUTE**
Enterprise Vocabulary Services

Search

National Cancer Institute

HomeAbout EVSToolsHistorySupportEVS Web Downloads

Related Links

[NCI Thesaurus](#)
[NCI Metathesaurus](#)
[NCI Term Browser](#)
[EVS Web Downloads](#)
[EVS FTP Downloads](#)
[FDA FTP Downloads](#)
[Vocabulary for Cancer Research](#)
[LexEVS Downloads](#)

NCI Thesaurus Downloads

The NCI is developed in [OWL](#) using NCI-developed plug-ins for the [Protege](#) editor, hence OWL is the primary distribution format and we provide asserted and inferred versions. We also provide the NCI in a tab-delimited format that some users prefer over XML for simple purposes.

In the files below, the **asserted** version contain is-a and role relations as stated by the editors of the NCI (except in the case of the tab-delimited file where role relations are excluded). The **inferred** version is the result of classification, and contains inferred is-a and role relations, as well as the asserted relations. In addition, annotations meant to support NCI editors, e.g. editor notes, are excluded from all files.

Please see the [NCI Thesaurus License](#) for terms of use.

Please see the [ReadMe.txt](#) file for more information on the download formats. Periodically it might contain additional information specific to a particular release version of the NCI.

NCI Thesaurus Downloads

Current Version: Editing of NCI Thesaurus 22.08e was completed on Aug 29, 2022. Version 22.08e was August's fifth build in the development cycle.

Thesaurus_22.08e.OWL.zip in OWL format (asserted version). This file contains the entire terminology expressed in OWL. The `rdf:about` (`rdf:ID`) utilizes semantically meaningless identifiers, e.g. "C123456", which is also the value of the "code" property. The `rdf:about` of annotation and object properties are also replaced with their corresponding codes. This is the version of the NCI that after classification (to generate an inferred version) is loaded into the LexEVS terminology server to support API and web browser access.

ThesaurusInf_22.08e.OWL.zip in OWL format (inferred version). This file is created for import into the UMLS and NCI Metathesaurus and excludes retired concepts. It uses semantically meaningless `rdf:about` identifiers as described above for the "asserted version".

Thesaurus_22.08e.FLAT.zip in tab-delimited format (asserted version). This file includes is-a relations, synonyms, and NCI definitions, but excludes roles and the rest of the annotation properties.

For previous releases of the NCI Thesaurus, browse the [Archives](#).

Home | Contact Us | Accessibility | Policies | HHS Vulnerability Disclosure | FOIA

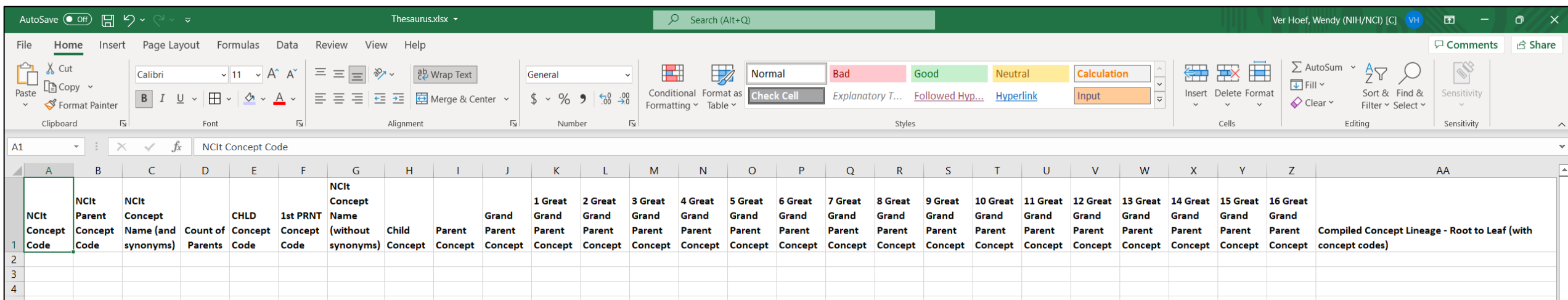
U.S. Department of Health and Human Services | National Institutes of Health | National Cancer Institute | USA.gov

Step 2: Unzip the file, extract the Thesaurus.txt file and open it using Excel as a tab-delimited file. Save the file as an .xlsx file.

AutoSave On							
Thesaurus.xlsx • Saved							
Search (Alt+Q)							
Ver Hoef, Wendy (NIH/NCI) [C]							
File Home Insert Page Layout Formulas Data Review View Help							
Clipboard Font Alignment Number Styles Cells Editing Sensitivity							
A1 C100000							
A	B	C	D	E	F	G	H
1	C100000	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100000>	C99521	Percutaneous Coronary Intervention for ST Elevation Myocardial Infarction-Stable Over	A percutaneous coronary intervention is necessary for a myocardial infarction that presents with ST segment elevation	Therapeutic or Preventive Procedure	
2	C100001	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100001>	C99521	Percutaneous Coronary Intervention for ST Elevation Myocardial Infarction-Stable After	A percutaneous coronary intervention is necessary for a myocardial infarction that presents with ST segment elevation	Therapeutic or Preventive Procedure	
3	C100002	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100002>	C99521	Percutaneous Coronary Intervention for ST Elevation Myocardial Infarction-Unstable	A percutaneous coronary intervention is necessary for a myocardial infarction that presents with ST segment elevation	Therapeutic or Preventive Procedure	
4	C100003	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100003>	C80449	Percutaneous Mitral Valve Repair PERCUTANEOUS MITRAL REPAIR	Invasive procedure through a puncture in the skin, performed for the purpose of reducing regurgitation of the mitral valve	Therapeutic or Preventive Procedure	
5	C100004	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100004>	C80430	Pericardial Stripping PERICARDIAL STRIPPING	Removal or resection of the pericardium. (ACC)	Therapeutic or Preventive Procedure	
6	C100005	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100005>	C15839	Post-Cardiac Transplant Evaluation POST-CARDIAC TRANSPLANT	A procedure to evaluate the health of the an individual after receiving a heart transplant. (ACC)	Health Care Activity	
7	C100006	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100006>	C139982	Pre-Operative Evaluation for Non-Cardiovascular Surgery PRE-OPERATIVE EVALUATION	A procedure to evaluate the health of the cardiac system to determine feasibility of a non-cardiac related surgical procedure	Health Care Activity	
8	C100007	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100007>	C99896	Previously Implanted Cardiac Lead PREVIOUSLY IMPLANTED	The coronary sinus / left ventricular lead was not implanted because a lead had been previously implanted. (ACC)	Finding	
9	C100008	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100008>	C99521	Rescue Percutaneous Coronary Intervention for ST Elevation Myocardial Infarction After	A percutaneous coronary intervention is imperative for a myocardial infarction that presents with ST segment elevation	Therapeutic or Preventive Procedure	
10	C100009	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100009>	C99896	Reused Cardiac Lead	The cardiac lead has been left in situ and reused. (ACC)	Finding	
11	C100010	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100010>	C61007	Cyclophosphamide/Fluoxymesterone/Mitotactol/Prednisone/Tamoxifen CD P TH CTX DBD/FXM/PRED/TMX	Obsolete_Concept	Therapeutic or Preventive Procedure	
12	C100010	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100010>	C80430	Right Cardiac Ventricle Assist Device Procedure	A procedure involving a mechanical pump which receives blood from the right ventricle and delivers it to the pulmonary	Therapeutic or Preventive Procedure	
13	C100011	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100011>	C63169	Right Cardiac Ventricle Pacing Malfunction RIGHT VENTRICULAR PACING MALFUNCTION	The cardiac rhythm device malfunction affected the right ventricular (RV) pacing. (ACC)	Event	
14	C100012	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100012>	C50797	Severe Cardiac Valve Regurgitation	Evidence of severe retrograde blood flow through the valve(s) of the heart. (ACC)	Disease or Syndrome	
15	C100013	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100013>	C62433	Severe Cardiac Valve Stenosis	The cardiac valve orifice is abnormally narrow, to a severe degree. (ACC)	Finding	
16	C100014	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100014>	C111103	Spontaneous Sustained Ventricular Tachycardia by ECG Finding SPONTANEOUS SUSTAINED	An electrocardiographic finding of ventricular tachycardia that occurs without provocation and does not resolve without	Finding	
17	C100015	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100015>	C99521	Staged Percutaneous Coronary Intervention STAGED PERCUTANEOUS CORONARY INTERVENTION	A subsequent Percutaneous Coronary Intervention in a planned series of procedures, where the first PCI could have been	Therapeutic or Preventive Procedure	
18	C100016	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100016>	C25476	Subject Delay in Providing Consent Patient Delay in Providing Consent SUBJECT DELAY	The subject has not provided consent in a timely manner for the procedure to take place. (ACC)	Temporal Concept	
19	C100017	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100017>	C80430	Surgical Maze Procedure Cox Maze SURGICAL MAZE	A cardiac procedure intended to inhibit atrial fibrillation, in which a series of incisions are made in the atria to block ablat	Therapeutic or Preventive Procedure	
20	C100018	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100018>	C35053	Syncope with High Risk Cardiac Characteristics SYNCOPE WITH HIGH RISK CHARACTERISTICS	A sudden loss of consciousness with loss of postural tone not related to anesthesia with high risk characteristics. High r	Finding	
21	C100019	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100019>	C111103	Syncope with Inducible Ventricular Tachycardia by ECG Finding SYNCOPE WITH INDUCIBLE	An electrocardiographic finding of ventricular tachycardia accompanied by syncope in the individual.	Finding	
22	C10001	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C10001>	C63510	CONPADRI-1 Regimen CONPADRI-1 COPADRI-1 CTX/DOX/L-PAM/VCR Cyclophosphamide	A chemotherapy regimen consisting of cyclophosphamide, vincristine, melphalan, and doxorubicin for the treatment of	Therapeutic or Preventive Procedure	
23	C100020	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100020>	C35317	Three Vessel Coronary Disease THREE VESSEL DISEASE	There was greater than or equal to 50% stenosis (reduction in cross-sectional area) in three coronary arteries (or greater	Disease or Syndrome	
24	C100021	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100021>	C35552	Thrombolysis in Myocardial Infarction Flow TIMI Flow TIMI Flow TIMIFLOW	A grading system for coronary blood flow based on the classification developed by the Thrombolysis in Myocardial Infar	Finding	
25	C100022	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100022>	C49803	Transcatheter Aortic Valve Implantation TRANSCATHETER AORTIC VALVE IMPLANTATION	Use of a thin tube to deliver an artificial valve apparatus through the left ventricle of the heart and the aorta. (ACC)	Therapeutic or Preventive Procedure	
26	C100023	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100023>	C35317	Two Vessel Coronary Disease TWO VESSEL DISEASE	There was greater than or equal to 50% stenosis (reduction in cross-sectional area) in two coronary arteries (or greater	Disease or Syndrome	
27	C100024	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100024>	C37254	Typical Coronary Artery Disease Symptom TYPICAL CORONARY ARTERY DISEASE SYMPTOM	Deep, poorly localized chest or arm discomfort that is reproducibly associated with physical exertion or emotional stress	Sign or Symptom	
28	C100025	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100025>	C99896	Unable to Obtain Acceptable Pacing Thresholds COULD NOT OBTAIN ACCEPTABLE PACING	The coronary sinus / left ventricular lead was not implanted because the operator was unable to obtain acceptable pac	Finding	
29	C100026	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100026>	C99896	Unable to Obtain Satisfactory Coronary Sinus Access UNABLE TO OBTAIN SATISFACTORY	The coronary sinus / left ventricular lead was not implanted because the operator was unable to obtain satisfactory co	Finding	
30	C100027	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100027>	C99896	Unable to Obtain Satisfactory Vascular Access UNABLE TO OBTAIN SATISFACTORY VASCULAR	The coronary sinus / left ventricular lead was not implanted because the operator was unable to obtain satisfactory va	Finding	
31	C100028	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100028>	C99896	Unable to Position Lead Due to Coronary Sinus Dissection UNABLE TO POSITION LEAD	The coronary sinus / left ventricular lead was not implanted because the coronary sinus was dissected. (ACC)	Finding	
32	C100029	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100029>	C99896	Unable to Position Lead In tributary Vein of Coronary Sinus UNABLE TO POSITION LEAD	The coronary sinus / left ventricular lead was not implanted because the operator was unable to position the lead in th	Finding	
33	C10002	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C10002>	C61063	Cyclophosphamide/Doxorubicin/Fluorouracil/Methotrexate/Vincristine CMP-V CTX/DOX/S-FU/MTX/PRED/VCR	Obsolete_Concept	Therapeutic or Preventive Procedure	
34	C100030	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100030>	C80430	Upgrade to Cardiac Device with Additional Pacing Capabilities UPGRADE TO A DEVICE	A device with additional pacing capabilities is replacing the current device. (ACC)	Therapeutic or Preventive Procedure	
35	C100031	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100031>	C49236	Urgent Procedure URGENT	Deferring an urgent procedure increases risk of adverse outcome or clinical deterioration. Clinical situation warrants p	Therapeutic or Preventive Procedure	
36	C100032	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100032>	C35552	American College of Cardiology/American Heart Association Lesion Complexity Class A	A classification system for coronary stenosis based upon characteristics that influence the difficulty of percutaneous c	Classification	
37	C100033	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100033>	C50479	Cardiac Arrest/Arrhythmia Etiology Unknown CARDIAC ARREST/ARRHYTHMIA ETIOLOGY	An episode of cardiac arrest or arrhythmia where the etiology was unknown. (ACC)	Finding	
38	C100034	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100034>	C99896	Cardiac Lead Dislodgement Requiring Reposition/Reoperation LEAD DISLODEMENT	The movement (macroscopic or microscopic) of an existing lead within the heart or vascular tree away from the origin	Finding	
39	C100035	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100035>	C62961	Faulty Device Connector/Header FAULTY CONNECTOR/HEADER	There is a faulty connector/header in the device. (ACC)	Event	
40	C100036	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100036>	C111103	Idiopathic Primary Ventricular Tachycardia/Ventricular Flutter by ECG Finding IDIOPATHIC	An electrocardiographic finding of ventricular tachycardia or ventricular fibrillation of unknown cause. (ACC)	Finding	
41	C100037	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100037>	C100032	American College of Cardiology/American Heart Association Lesion Complexity Score B	Discrete (less than 10 mm length) and concentric and readily accessible and non-angled segment less than 45 degree	Finding	
42	C100038	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100038>	C100032	American College of Cardiology/American Heart Association Lesion Complexity Score B	Tubular (10-20 mm length) or eccentric or moderate tortuosity of proximal segment or moderately angulated segment	Finding	
43	C100039	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C100039>	C100032	American College of Cardiology/American Heart Association Lesion Complexity Score C	Diffuse (length greater than 2cm) or excessive tortuosity of proximal segment or extremely angulated segments greater	Finding	
44	C10003	<http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C10003>	C61063	Methotrexate/Teniposide MTX/VM-26	Obsolete_Concept	Therapeutic or Preventive Procedure	

Step 3: Add a new tab called “NCIt Lineage Lookup” with the following column headings:

- | | | | |
|----|--------------------------------------|-----|--|
| A. | NCIt Concept Code | O. | 5 Great Grand Parent Concept |
| B. | NCIt Parent Concept Code | P. | 6 Great Grand Parent Concept |
| C. | NCIt Concept Name (and synonyms) | Q. | 7 Great Grand Parent Concept |
| D. | Count of Parents | R. | 8 Great Grand Parent Concept |
| E. | CHLD Concept Code | S. | 9 Great Grand Parent Concept |
| F. | 1st PRNT Concept Code | T. | 10 Great Grand Parent Concept |
| G. | NCIt Concept Name (without synonyms) | U. | 11 Great Grand Parent Concept |
| H. | Child Concept | V. | 12 Great Grand Parent Concept |
| I. | Parent Concept | W. | 13 Great Grand Parent Concept |
| J. | Grand Parent Concept | X. | 14 Great Grand Parent Concept |
| K. | 1 Great Grand Parent Concept | Y. | 15 Great Grand Parent Concept |
| L. | 2 Great Grand Parent Concept | Z. | 16 Great Grand Parent Concept |
| M. | 3 Great Grand Parent Concept | AA. | Compiled Concept Lineage - Root to Leaf (with concept codes) |
| N. | 4 Great Grand Parent Concept | | |



BEFORE Step 4: A word about some concepts having multiple parents & synonyms

- Numerous concepts have more than one parent, separated by “|”
 - For the sake of simplicity, and since there is no indication of priority, this process uses only the first parent to trace lineage, though the other parents’ lineages may be just as valid and useful
- Numerous concepts have one or more synonyms, separated by “|”
 - For the purposes of identifying the lineage, synonyms are not particularly useful, so are truncated from the preferred term by this process
- A few examples follow...

Concept ID	Parent Concept ID(s)	Concept Term [Synonym(s)]
C100063	C3824 C35552	Coronary Vessel Lesion
C100068	C80430 C20985	Cardiac Ablation Catheter ablation
C100093	C5461 C121619	Central Nervous System Nongerminomatous Germ Cell Tumor Nongerminomatous GCT - CNS Nongerminomatous germ cell tumor of the CNS
C101132	C2201 C1742 C129825	mTOR1/2 Kinase Inhibitor ME-344 ME-344 ME-344
C101257	C129826 C129825	IGF-1R Antisense Oligodeoxynucleotide-treated Autologous Glioma Cells IGV-001 IGF-1R/AS ODN-treated Autologous Glioma Cells IGV-001 IGV 001 IGV-001 IGV001
C101258	C28676 C137999 C129826	CD19CAR-CD3zeta-4-1BB-expressing Allogeneic T-lymphocyte Cells CART-19 Cells
C101786	C61074 C129825	Multi-AGC Kinase Inhibitor AT13148 AT-13148 AT13148
C101787	C28681 C138180 C129826	TIL 1383I T Cell Receptor-Transduced Autologous T Cells TIL 1383I TCR Transduced Autologous T Cells

Step 4: Copy Formulas into the “NCIt Lineage Lookup” tab & propagate down to 250K rows

- Current downloads contain >175K rows, so propagating the formulas down to row 250,001 will cover the needs for the near-term future
- First 3 columns copy the concept code, parent concept code and concept name & synonyms from the Thesaurus tab
- Thereafter all columns are derived from the first 3
- Functions are used to remove the synonyms from the preferred term, look up child concepts to identify the parent concepts, and then look up and concatenate the series of concept names for each of the generations in the lineage of the concept
- Conditional functions are used to ignore columns when the generations in the lineage of any given concept have reached the root concept
- Most formulas involve hardcoded references to the range of child & parent concept IDs using the “\$” symbol which prevents column letters and row numbers from automatically increasing and you drag the cells to propagate the formulas
 - For example, “\$E\$2:\$F\$250001” which references the cells between column E, row 2, and column F, row 250001 for child-to-parent lookup
- The list of formulas on next slide...

Column	Formula
NCIt Concept Code	Copy values from Thesaurus tab, column A
NCIt Parent Concept Code	Copy values from Thesaurus tab, column C
NCIt Concept Name (and synonyms)	Copy values from Thesaurus tab, column D
Count of Parents	=LEN(B2)-LEN(SUBSTITUTE(B2," ", ""))+1
CHLD Concept Code	="CHLD-"&A2
1st PRNT Concept Code	=IF(D2=1,"PRNT-"&TRIM(B2),"PRNT-"&MID(TRIM(B2),1,FIND(" ",TRIM(B2))-1))
NCIt Concept Name (without synonyms)	=IF((LEN(C2)-LEN(SUBSTITUTE(C2," ", ""))+1)=1,TRIM(C2),MID(TRIM(C2),1,FIND(" ",TRIM(C2))-1))
Child Concept	=SUBSTITUTE(E2,"CHLD-", "")
Parent Concept	=SUBSTITUTE(F2,"PRNT-", "")
Grand Parent Concept	=IF(LEN(I2)=0, "",SUBSTITUTE(VLOOKUP("CHLD-"&I2,\$E\$2:\$F\$250001,2,FALSE),"PRNT-", ""))
1 Great Grand Parent Concept	=IF(LEN(J2)=0, "",SUBSTITUTE(VLOOKUP("CHLD-"&J2,\$E\$2:\$F\$250001,2,FALSE),"PRNT-", ""))
2 Great Grand Parent Concept	=IF(LEN(K2)=0, "",SUBSTITUTE(VLOOKUP("CHLD-"&K2,\$E\$2:\$F\$250001,2,FALSE),"PRNT-", ""))
3 Great Grand Parent Concept	=IF(LEN(L2)=0, "",SUBSTITUTE(VLOOKUP("CHLD-"&L2,\$E\$2:\$F\$250001,2,FALSE),"PRNT-", ""))
4 Great Grand Parent Concept	=IF(LEN(M2)=0, "",SUBSTITUTE(VLOOKUP("CHLD-"&M2,\$E\$2:\$F\$250001,2,FALSE),"PRNT-", ""))
5 Great Grand Parent Concept	=IF(LEN(N2)=0, "",SUBSTITUTE(VLOOKUP("CHLD-"&N2,\$E\$2:\$F\$250001,2,FALSE),"PRNT-", ""))
6 Great Grand Parent Concept	=IF(LEN(O2)=0, "",SUBSTITUTE(VLOOKUP("CHLD-"&O2,\$E\$2:\$F\$250001,2,FALSE),"PRNT-", ""))
7 Great Grand Parent Concept	=IF(LEN(P2)=0, "",SUBSTITUTE(VLOOKUP("CHLD-"&P2,\$E\$2:\$F\$250001,2,FALSE),"PRNT-", ""))
8 Great Grand Parent Concept	=IF(LEN(Q2)=0, "",SUBSTITUTE(VLOOKUP("CHLD-"&Q2,\$E\$2:\$F\$250001,2,FALSE),"PRNT-", ""))
9 Great Grand Parent Concept	=IF(LEN(R2)=0, "",SUBSTITUTE(VLOOKUP("CHLD-"&R2,\$E\$2:\$F\$250001,2,FALSE),"PRNT-", ""))
10 Great Grand Parent Concept	=IF(LEN(S2)=0, "",SUBSTITUTE(VLOOKUP("CHLD-"&S2,\$E\$2:\$F\$250001,2,FALSE),"PRNT-", ""))
11 Great Grand Parent Concept	=IF(LEN(T2)=0, "",SUBSTITUTE(VLOOKUP("CHLD-"&T2,\$E\$2:\$F\$250001,2,FALSE),"PRNT-", ""))
12 Great Grand Parent Concept	=IF(LEN(U2)=0, "",SUBSTITUTE(VLOOKUP("CHLD-"&U2,\$E\$2:\$F\$250001,2,FALSE),"PRNT-", ""))
13 Great Grand Parent Concept	=IF(LEN(V2)=0, "",SUBSTITUTE(VLOOKUP("CHLD-"&V2,\$E\$2:\$F\$250001,2,FALSE),"PRNT-", ""))
14 Great Grand Parent Concept	=IF(LEN(W2)=0, "",SUBSTITUTE(VLOOKUP("CHLD-"&W2,\$E\$2:\$F\$250001,2,FALSE),"PRNT-", ""))
15 Great Grand Parent Concept	=IF(LEN(X2)=0, "",SUBSTITUTE(VLOOKUP("CHLD-"&X2,\$E\$2:\$F\$250001,2,FALSE),"PRNT-", ""))
16 Great Grand Parent Concept	=IF(LEN(Y2)=0, "",SUBSTITUTE(VLOOKUP("CHLD-"&Y2,\$E\$2:\$F\$250001,2,FALSE),"PRNT-", ""))
Compiled Concept Lineage - Root to Leaf (with concept codes)	=IF(LEN(Z2)=0, "",VLOOKUP("CHLD-"&Z2,\$E\$2:\$G\$250001,3,FALSE)&" ("&Z2&") > ")&IF(LEN(Y2)=0, "",VLOOKUP("CHLD-"&Y2,\$E\$2:\$G\$250001,3,FALSE)&" ("&Y2&") > ")&IF(LEN(X2)=0, "",VLOOKUP("CHLD-"&X2,\$E\$2:\$G\$250001,3,FALSE)&" ("&X2&") > ")&IF(LEN(W2)=0, "",VLOOKUP("CHLD-"&W2,\$E\$2:\$G\$250001,3,FALSE)&" ("&W2&") > ")&IF(LEN(V2)=0, "",VLOOKUP("CHLD-"&V2,\$E\$2:\$G\$250001,3,FALSE)&" ("&V2&") > ")&IF(LEN(U2)=0, "",VLOOKUP("CHLD-"&U2,\$E\$2:\$G\$250001,3,FALSE)&" ("&U2&") > ")&IF(LEN(T2)=0, "",VLOOKUP("CHLD-"&T2,\$E\$2:\$G\$250001,3,FALSE)&" ("&T2&") > ")&IF(LEN(S2)=0, "",VLOOKUP("CHLD-"&S2,\$E\$2:\$G\$250001,3,FALSE)&" ("&S2&") > ")&IF(LEN(R2)=0, "",VLOOKUP("CHLD-"&R2,\$E\$2:\$G\$250001,3,FALSE)&" ("&R2&") > ")&IF(LEN(Q2)=0, "",VLOOKUP("CHLD-"&Q2,\$E\$2:\$G\$250001,3,FALSE)&" ("&Q2&") > ")&IF(LEN(P2)=0, "",VLOOKUP("CHLD-"&P2,\$E\$2:\$G\$250001,3,FALSE)&" ("&P2&") > ")&IF(LEN(O2)=0, "",VLOOKUP("CHLD-"&O2,\$E\$2:\$G\$250001,3,FALSE)&" ("&O2&") > ")&IF(LEN(N2)=0, "",VLOOKUP("CHLD-"&N2,\$E\$2:\$G\$250001,3,FALSE)&" ("&N2&") > ")&IF(LEN(M2)=0, "",VLOOKUP("CHLD-"&M2,\$E\$2:\$G\$250001,3,FALSE)&" ("&M2&") > ")&IF(LEN(L2)=0, "",VLOOKUP("CHLD-"&L2,\$E\$2:\$G\$250001,3,FALSE)&" ("&L2&") > ")&IF(LEN(K2)=0, "",VLOOKUP("CHLD-"&K2,\$E\$2:\$G\$250001,3,FALSE)&" ("&K2&") > ")&IF(LEN(J2)=0, "",VLOOKUP("CHLD-"&J2,\$E\$2:\$G\$250001,3,FALSE)&" ("&J2&") > ")&IF(LEN(I2)=0, "",VLOOKUP("CHLD-"&I2,\$E\$2:\$G\$250001,3,FALSE)&" ("&I2&") > ")&VLOOKUP("CHLD-"&H2,\$E\$2:\$G\$250001,3,FALSE)&" ("&H2&") "

Copy formulas and paste on the NCIt Lineage Lookup tab. Then select the cells in the new row and “drag” them down to through 250,000 rows (or at least as many rows as are on the Thesaurus tab).

Results of Step 4 look like this...

Since each child can be the parent of multiple other concepts, prefixing the child concept makes it unique within the list

This is a root concept, so no more values in great grandparent columns to the right

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
	NCIt Concept Code	NCIt Parent Concept Code	NCIt Concept Name (and synonyms)	Count of Parents	CHLD Concept Code	1st PRNT Concept Code	NCIt Concept Name (without synonyms)	Child Concept	Parent Concept	Grand Parent Concept	1 Great Grand Parent Concept	2 Great Grand Parent Concept	3 Great Grand Parent Concept	4 Great Grand Parent Concept	5 Great Grand Parent Concept	6 Great Grand Parent Concept	7 Great Grand Parent Concept	8 Great Grand Parent Concept
1	C100000	C99521	Percutaneous Coronary Intervention for ST Elevation Myocardial Infarction-Stable	1	CHLD-C100000	PRNT-C99521	Percutaneous Coronary Intervention for ST Elevation Myocardial Infarction-Stable	C100000	C99521	C49803	C49163	C15329	C25218	C16203	C43431			
2	C100001	C99521	Percutaneous Coronary Intervention for ST Elevation Myocardial Infarction-Stable	1	CHLD-C100001	PRNT-C99521	Percutaneous Coronary Intervention for ST Elevation Myocardial Infarction-Stable	C100001	C99521	C49803	C49163	C15329	C25218	C16203	C43431			
3	C100002	C99521	Percutaneous Coronary Intervention for ST Elevation Myocardial Infarction-Unstable	1	CHLD-C100002	PRNT-C99521	Percutaneous Coronary Intervention for ST Elevation Myocardial Infarction-Unstable	C100002	C99521	C49803	C49163	C15329	C25218	C16203	C43431			
4	C100003	C80449	Percutaneous Mitral Valve Repair PERCUTANEOUS MITRAL REPAIR	1	CHLD-C100003	PRNT-C80449	Percutaneous Mitral Valve Repair	C100003	C80449	C50818	C99546	C80430	C49236	C25218	C16203	C43431		
5	C100004	C80430	Pericardial Stripping PERICARDIAL STRIPPING	1	CHLD-C100004	PRNT-C80430	Pericardial Stripping	C100004	C80430	C49236	C25218	C16203	C43431					
6	C100005	C15839	Post-Cardiac Transplant Evaluation POST-CARDIAC TRANSPLANT	1	CHLD-C100005	PRNT-C15839	Post-Cardiac Transplant Evaluation	C100005	C15839	C15857	C15791	C15336	C15429	C16203	C43431			
7	C100006	C139982	Pre-Operative Evaluation for Non-Cardiovascular Surgery PRE-OPERATIVE EVALUATION	1	CHLD-C100006	PRNT-C139982	Pre-Operative Evaluation for Non-Cardiovascular Surgery	C100006	C139982	C15839	C15857	C15791	C15336	C15429	C16203	C43431		
8	C100007	C99896	Previously Implanted Cardiac Lead PREVIOUSLY IMPLANTED	1	CHLD-C100007	PRNT-C99896	Previously Implanted Cardiac Lead	C100007	C99896	C35552	C36278	C3367	C7057					
9	C100008	C99521	Rescue Percutaneous Coronary Intervention for ST Elevation Myocardial Infarction	1	CHLD-C100008	PRNT-C99521	Rescue Percutaneous Coronary Intervention for ST Elevation Myocardial Infarction	C100008	C99521	C49803	C49163	C15329	C25218	C16203	C43431			
10	C100009	C99896	Reused Cardiac Lead	1	CHLD-C100009	PRNT-C99896	Reused Cardiac Lead	C100009	C99896	C35552	C36278	C3367	C7057					
11	C100010	C61007	Cyclophosphamide/Fluoxymesterone/Mitotactol/Prednisone/Tamoxifen CD(P)TH	1	CHLD-C100010	PRNT-C61007	Cyclophosphamide/Fluoxymesterone/Mitotactol/Prednisone/Tamoxifen CD(P)TH	C100010	C61007	C62633	C12218							
12	C100011	C80430	Right Cardiac Ventricle Assist Device Procedure	1	CHLD-C100011	PRNT-C80430	Right Cardiac Ventricle Assist Device Procedure	C100011	C80430	C49236	C25218	C16203	C43431					
13	C100012	C63169	Right Cardiac Ventricle Pacing Malfunction RIGHT VENTRICULAR PACING MALFUNCTION	1	CHLD-C100012	PRNT-C63169	Right Cardiac Ventricle Pacing Malfunction	C100012	C63169	C63007	C62954	C62996	C54034	C20189				
14	C100013	C50797	Severe Cardiac Valve Regurgitation	1	CHLD-C100013	PRNT-C50797	Severe Cardiac Valve Regurgitation	C100013	C50797	C45525	C35741	C35743	C53531	C53529	C2991	C7057		
15	C100014	C62433	Severe Cardiac Valve Stenosis	1	CHLD-C100014	PRNT-C62433	Severe Cardiac Valve Stenosis	C100014	C62433	C50754	C36287	C53547	C53529	C2991	C7057			
16	C100015	C111103	Spontaneous Sustained Ventricular Tachycardia by ECG Finding SPONTANEOUS SUSTAINED VENTRICULAR TACHYCARDIA BY ECG FINDING	1	CHLD-C100015	PRNT-C111103	Spontaneous Sustained Ventricular Tachycardia by ECG Finding	C100015	C111103	C111101	C111075	C41328	C35552	C36278	C3367	C7057		
17	C100016	C99521	Staged Percutaneous Coronary Intervention STAGED PERCUTANEOUS CORONARY INTERVENTION	1	CHLD-C100016	PRNT-C99521	Staged Percutaneous Coronary Intervention	C100016	C99521	C49803	C49163	C15329	C25218	C16203	C43431			
18	C100017	C25476	Subject Delay in Providing Consent Patient Delay in Providing Consent SUBJECT DELAY IN PROVIDING CONSENT	1	CHLD-C100017	PRNT-C25476	Subject Delay in Providing Consent	C100017	C25476	C21514	C41009	C20189						
19	C100018	C80430	Surgical Maze Procedure Surgical Maze Procedure	1	CHLD-C100018	PRNT-C80430	Surgical Maze Procedure	C100018	C80430	C49236	C25218	C16203	C43431					

The Lineage starts at the root concept & goes through all generations to the leaf node concept

	S	T	U	V	W	X	Y	Z	AA
	9 Great Grand Parent Concept	10 Great Grand Parent Concept	11 Great Grand Parent Concept	12 Great Grand Parent Concept	13 Great Grand Parent Concept	14 Great Grand Parent Concept	15 Great Grand Parent Concept	16 Great Grand Parent Concept	
1									Compiled Concept Lineage - Root to Leaf (with concept codes)
2									Activity (C43431) > Clinical or Research Activity (C16203) > Intervention or Procedure (C25218) > Surgical Procedure (C15329) > Surgical Procedure by Site or System (C49163) > Cardiovascular Surgical Procedure (C49803) > Percutaneous Coronary Intervention
3									Activity (C43431) > Clinical or Research Activity (C16203) > Intervention or Procedure (C25218) > Surgical Procedure (C15329) > Surgical Procedure by Site or System (C49163) > Cardiovascular Surgical Procedure (C49803) > Percutaneous Coronary Intervention
4									Activity (C43431) > Clinical or Research Activity (C16203) > Intervention or Procedure (C25218) > Surgical Procedure (C15329) > Surgical Procedure by Site or System (C49163) > Cardiovascular Surgical Procedure (C49803) > Percutaneous Coronary Intervention
5									Activity (C43431) > Clinical or Research Activity (C16203) > Intervention or Procedure (C25218) > Therapeutic Procedure (C49236) > Cardiac Therapeutic Procedure (C80430) > Cardiac Therapeutic Procedure (C80430) > Right Cardiac Ventricle Assist Device Procedure (C100010)
6									Activity (C43431) > Clinical or Research Activity (C16203) > Intervention or Procedure (C25218) > Therapeutic Procedure (C49236) > Cardiac Therapeutic Procedure (C80430) > Cardiac Therapeutic Procedure (C80430) > Right Cardiac Ventricle Assist Device Procedure (C100010)
7									Activity (C43431) > Clinical or Research Activity (C16203) > Research Activity (C15429) > Testing (C15336) > Clinical Testing (C15791) > Therapy Evaluation (C15857) > Prevention and Treatment Evaluation (C15839) > Pre-operative Evaluation (C139982) > Pre-operative Evaluation (C139982) > Pre-operative Evaluation (C139982)
8									Activity (C43431) > Clinical or Research Activity (C16203) > Research Activity (C15429) > Testing (C15336) > Clinical Testing (C15791) > Therapy Evaluation (C15857) > Prevention and Treatment Evaluation (C15839) > Pre-operative Evaluation (C139982) > Pre-operative Evaluation (C139982) > Pre-operative Evaluation (C139982)
9									Disease, Disorder or Finding (C7057) > Finding (C3367) > Finding by Site or System (C36278) > Cardiovascular System Finding (C35552) > Cardiac Lead Event Finding (C99896) > Previously Implanted Cardiac Lead (C100007)
10									Activity (C43431) > Clinical or Research Activity (C16203) > Intervention or Procedure (C25218) > Surgical Procedure (C15329) > Surgical Procedure by Site or System (C49163) > Cardiovascular Surgical Procedure (C49803) > Percutaneous Coronary Intervention
11									Disease, Disorder or Finding (C7057) > Finding (C3367) > Finding by Site or System (C36278) > Cardiovascular System Finding (C35552) > Cardiac Lead Event Finding (C99896) > Reused Cardiac Lead (C100009)
12									Chemotherapy Regimen or Agent Combination (C12218) > Agent Combination Indexed in Open Clinical Trials (C61007) > Cyclophosphamide/Fluoxymesterone/Mitotactol/Prednisone/Tamoxifen (C10000)
13									Activity (C43431) > Clinical or Research Activity (C16203) > Intervention or Procedure (C25218) > Therapeutic Procedure (C49236) > Cardiac Therapeutic Procedure (C80430) > Right Cardiac Ventricle Assist Device Procedure (C100010)
14									Property or Attribute (C20189) > Reportable Event (C54034) > Medical Device Failure or Malfunction (C62996) > Medical Device Operational Issue (C62954) > Electrical/Electronic Property Problem (C63007) > Pacing Problem (C63169) > Right Cardiac Ventricle Pacing Malfunction (C63169)
15									Disease, Disorder or Finding (C7057) > Disease or Disorder (C2991) > Non-Neoplastic Disorder (C53529) > Non-Neoplastic Disorder by Site (C53531) > Non-Neoplastic Thoracic Disorder (C35743) > Non-Neoplastic Heart Disorder (C35741) > Valvular Heart Disorder (C35741)
16									Disease, Disorder or Finding (C7057) > Disease or Disorder (C2991) > Non-Neoplastic Disorder (C53529) > Non-Neoplastic Disorder by Special Category (C53547) > Congenital or Acquired Anatomic Abnormality (C36287) > Stenosis (C50754) > Valvular Stenosis (C50754)
17									Disease, Disorder or Finding (C7057) > Finding (C3367) > Finding by Site or System (C36278) > Cardiovascular System Finding (C35552) > Electrocardiographic Finding (C41328) > Arrhythmia by ECG Finding (C111075) > Ventricular Arrhythmia by ECG Finding (C111075)
18									Activity (C43431) > Clinical or Research Activity (C16203) > Intervention or Procedure (C25218) > Surgical Procedure (C15329) > Surgical Procedure by Site or System (C49163) > Cardiovascular Surgical Procedure (C49803) > Percutaneous Coronary Intervention
19									Property or Attribute (C20189) > Qualifier (C41009) > Temporal Qualifier (C21514) > Delay (C25476) > Subject Delay in Providing Consent (C100016)
20									Activity (C43431) > Clinical or Research Activity (C16203) > Intervention or Procedure (C25218) > Therapeutic Procedure (C49236) > Cardiac Therapeutic Procedure (C80430) > Surgical Maze Procedure (C100017)

Step 5: Create a new tab, “Concept Code and Lineage”, copy & paste 1st & last columns as text

	A	B
	NCIt Concept Code	
1	Compiled Concept Lineage - Root to Leaf (with concept codes)	
2	C100000	Activity (C43431) > Clinical or Research Activity (C16203) > Intervention or Procedure (C25218) > Surgical Procedure (C15329) > Surgical Procedure by Site or System (C49163) > Cardiovascular Surgical Procedure (C49803) > Percutaneous Coronary Intervention (C99521) > Percutaneous Coronary Intervention for ST Elevation
3	C100001	Activity (C43431) > Clinical or Research Activity (C16203) > Intervention or Procedure (C25218) > Surgical Procedure (C15329) > Surgical Procedure by Site or System (C49163) > Cardiovascular Surgical Procedure (C49803) > Percutaneous Coronary Intervention (C99521) > Percutaneous Coronary Intervention for ST Elevation
4	C100002	Activity (C43431) > Clinical or Research Activity (C16203) > Intervention or Procedure (C25218) > Surgical Procedure (C15329) > Surgical Procedure by Site or System (C49163) > Cardiovascular Surgical Procedure (C49803) > Percutaneous Coronary Intervention (C99521) > Percutaneous Coronary Intervention for ST Elevation
5	C100003	Activity (C43431) > Clinical or Research Activity (C16203) > Intervention or Procedure (C25218) > Therapeutic Procedure (C49236) > Cardiac Therapeutic Procedure (C80430) > Cardiac Valve Procedure (C99546) > Heart Valve Repair (C50818) > Mitral Valve Repair (C80449) > Percutaneous Mitral Valve Repair (C100003)
6	C100004	Activity (C43431) > Clinical or Research Activity (C16203) > Intervention or Procedure (C25218) > Therapeutic Procedure (C49236) > Cardiac Therapeutic Procedure (C80430) > Pericardial Stripping (C100004)
7	C100005	Activity (C43431) > Clinical or Research Activity (C16203) > Research Activity (C15429) > Testing (C15336) > Clinical Testing (C15791) > Therapy Evaluation (C15857) > Prevention and Treatment Evaluation (C15839) > Post-Cardiac Transplant Evaluation (C100005)
8	C100006	Activity (C43431) > Clinical or Research Activity (C16203) > Research Activity (C15429) > Testing (C15336) > Clinical Testing (C15791) > Therapy Evaluation (C15857) > Prevention and Treatment Evaluation (C15839) > Pre-operative Evaluation (C139982) > Pre-Operative Evaluation for Non-Cardiovascular Surgery (C100006)
9	C100007	Disease, Disorder or Finding (C7057) > Finding (C3367) > Finding by Site or System (C36278) > Cardiovascular System Finding (C35552) > Cardiac Lead Event Finding (C99896) > Previously Implanted Cardiac Lead (C100007)
10	C100008	Activity (C43431) > Clinical or Research Activity (C16203) > Intervention or Procedure (C25218) > Surgical Procedure (C15329) > Surgical Procedure by Site or System (C49163) > Cardiovascular Surgical Procedure (C49803) > Percutaneous Coronary Intervention (C99521) > Rescue Percutaneous Coronary Intervention for ST Elevation
11	C100009	Disease, Disorder or Finding (C7057) > Finding (C3367) > Finding by Site or System (C36278) > Cardiovascular System Finding (C35552) > Cardiac Lead Event Finding (C99896) > Reused Cardiac Lead (C100009)
12	C10000	Chemotherapy Regimen or Agent Combination (C12218) > Agent Combination (C62633) > Agent Combination Indexed in Open Clinical Trials (C61007) > Cyclophosphamide/Fluoxymesterone/Mitoxantrone/Prednisone/Tamoxifen (C10000)
13	C100010	Activity (C43431) > Clinical or Research Activity (C16203) > Intervention or Procedure (C25218) > Therapeutic Procedure (C49236) > Cardiac Therapeutic Procedure (C80430) > Right Cardiac Ventricle Assist Device Procedure (C100010)
14	C100011	Property or Attribute (C20189) > Reportable Event (C54034) > Medical Device Failure or Malfunction (C62996) > Medical Device Operational Issue (C62954) > Electrical/Electronic Property Problem (C63007) > Pacing Problem (C63169) > Right Cardiac Ventricle Pacing Malfunction (C100011)
15	C100012	Disease, Disorder or Finding (C7057) > Disease or Disorder (C2991) > Non-Neoplastic Disorder (C53529) > Non-Neoplastic Disorder by Site (C53531) > Non-Neoplastic Thoracic Disorder (C35743) > Non-Neoplastic Heart Disorder (C45525) > Valvular Heart Disorder (C45525) > Valvular Regurgitation (C50797) > Severe Cardiac
16	C100013	Disease, Disorder or Finding (C7057) > Disease or Disorder (C2991) > Non-Neoplastic Disorder (C53529) > Non-Neoplastic Disorder by Special Category (C53547) > Congenital or Acquired Anatomic Abnormality (C36287) > Stenosis (C50754) > Valvular Stenosis (C62433) > Severe Cardiac Valve Stenosis (C100013)
17	C100014	Disease, Disorder or Finding (C7057) > Finding (C3367) > Finding by Site or System (C36278) > Cardiovascular System Finding (C35552) > Electrocardiographic Finding (C41328) > Arrhythmia by ECG Finding (C111075) > Ventricular Arrhythmia by ECG Finding (C111101) > Ventricular Tachycardia by ECG Finding (C111103) > Spont
18	C100015	Activity (C43431) > Clinical or Research Activity (C16203) > Intervention or Procedure (C25218) > Surgical Procedure (C15329) > Surgical Procedure by Site or System (C49163) > Cardiovascular Surgical Procedure (C49803) > Percutaneous Coronary Intervention (C99521) > Staged Percutaneous Coronary Intervention (C100015)
19	C100016	Property or Attribute (C20189) > Qualifier (C41009) > Temporal Qualifier (C21514) > Delay (C25476) > Subject Delay in Providing Consent (C100016)
20	C100017	Activity (C43431) > Clinical or Research Activity (C16203) > Intervention or Procedure (C25218) > Therapeutic Procedure (C49236) > Cardiac Therapeutic Procedure (C80430) > Surgical Maze Procedure (C100017)

- Copying the first and last columns from the formulas tab and pasting as text on the new tab provides an artifact that can be copied by itself and pasted elsewhere without dependency on the formulas
- This new tab can be used as a lookup to obtain the lineage of any concept in NCIt
- Note that it may be out of date the next time NCIt publishes a release



Compressed
(zipped) Folder

Example file –
“Thesaurus
(downloaded
20220921).zip”

How to use the NCIt Concept Lineage lookup

- Once the concept code and lineage have been copied onto a tab in another file, any C-code can be looked up using a formula

`=IF(ISBLANK(E2), "", VLOOKUP(TRIM(E2), 'NCIt Concept Code and Lineage'!A2:'NCIt Concept Code and Lineage'!B174127, 2, FALSE))`

1.) The value being looked up is in cell E2. If it is blank, the function returns an empty string.

2.) Otherwise, trim trailing blanks off the value being looked up.

3.) Look up the value on the "NCIt Concept Code and Lineage" tab, in the cells ranging from cell A2 to B174127.

4.) The "2" indicates that it's the 2nd column (B) whose value to return from the lookup. The "FALSE" indicates it should be an exact match, not approximate.

The screenshot shows an Excel spreadsheet with the following data table:

	A	B	C	D	E
1	Domain Name	CDE Name	NCIt Concept Lineage	NCIt concept name	NCIt CUI
2	Demographics	Unique ID	Conceptual Entity (C20181) > Intellectual Property (C97331) > Document (C19498) > Identifier (C25364) > Unique Identifier (C70663)	Unique Identifier	C70663
3	Demographics	Full Name	Property or Attribute (C20189) > Qualifier (C41009) > General Qualifier (C27993) > Full (C25517)	Full	C25517
4	Demographics	Full Name	Conceptual Entity (C20181) > Name (C42614) > Person Name (C25191)	Person Name	C25191
5	Demographics	Last Name	Conceptual Entity (C20181) > Name (C42614) > Person Name (C25191) > Last Name (C40975)	Last Name	C40975
6	Demographics	Age	Property or Attribute (C20189) > Person/Individual Attribute (C171087) > Personal Attribute (C19332) > Age (C25150)	Age	C25150
7	Demographics	Age Units	Property or Attribute (C20189) > Person/Individual Attribute (C171087) > Personal Attribute (C19332) > Age (C25150)	Age	C25150
8	Demographics	Age Units	Property or Attribute (C20189) > Unit of Measure (C25709)	Unit of Measure	C25709

The formula bar shows the formula: `=IF(ISBLANK(E2), "", VLOOKUP(TRIM(E2), 'NCIt Concept Code and Lineage'!A2:'NCIt Concept Code and Lineage'!B174127, 2, FALSE))`. A red arrow points from the explanation text to the 'NCIt CUI' cell (E2) in the table.