



Sander Su

FDA Orphan Drug Data Analysis

04-27-2021

Background



andrewsu commented 5 days ago

Member ...

FDA maintains a database of Orphan Drug Designations and Approvals (which currently has 5851 entries) at <https://www.accessdata.fda.gov/scripts/opdlisting/oopd/>. Some cleaning and normalization of both diseases and drugs will be required, so perhaps not the simplest data source to ingest...

To download HTML table version of results, select "Download Excel file" option (yes, it's mislabeled, but Excel will read it in):

The screenshot shows the FDA's website for searching orphan drug designations. The top navigation bar includes links for Home, Food, Drugs, Medical Devices, Radiation-Emitting Products, Vaccines, Blood & Biologics, Animal & Veterinary, Cosmetics, and Tobacco Products. The main content area is titled "Search Orphan Drug Designations and Approvals". It features a search form with fields for Product Name, Orphan Designation, Start Date (01/01/1983), End Date (04/09/2021), and Search results (All designations). A dropdown menu for "Output format" is open, showing "Download Excel file" as the selected option. Other options in the dropdown include "Display condensed list" and "Display detailed list". A note below the dropdown states: "Note: The information contained in the downloadable Excel file for the 'CF Grid Key' column is for FDA purposes only."

Data Set (1/2)

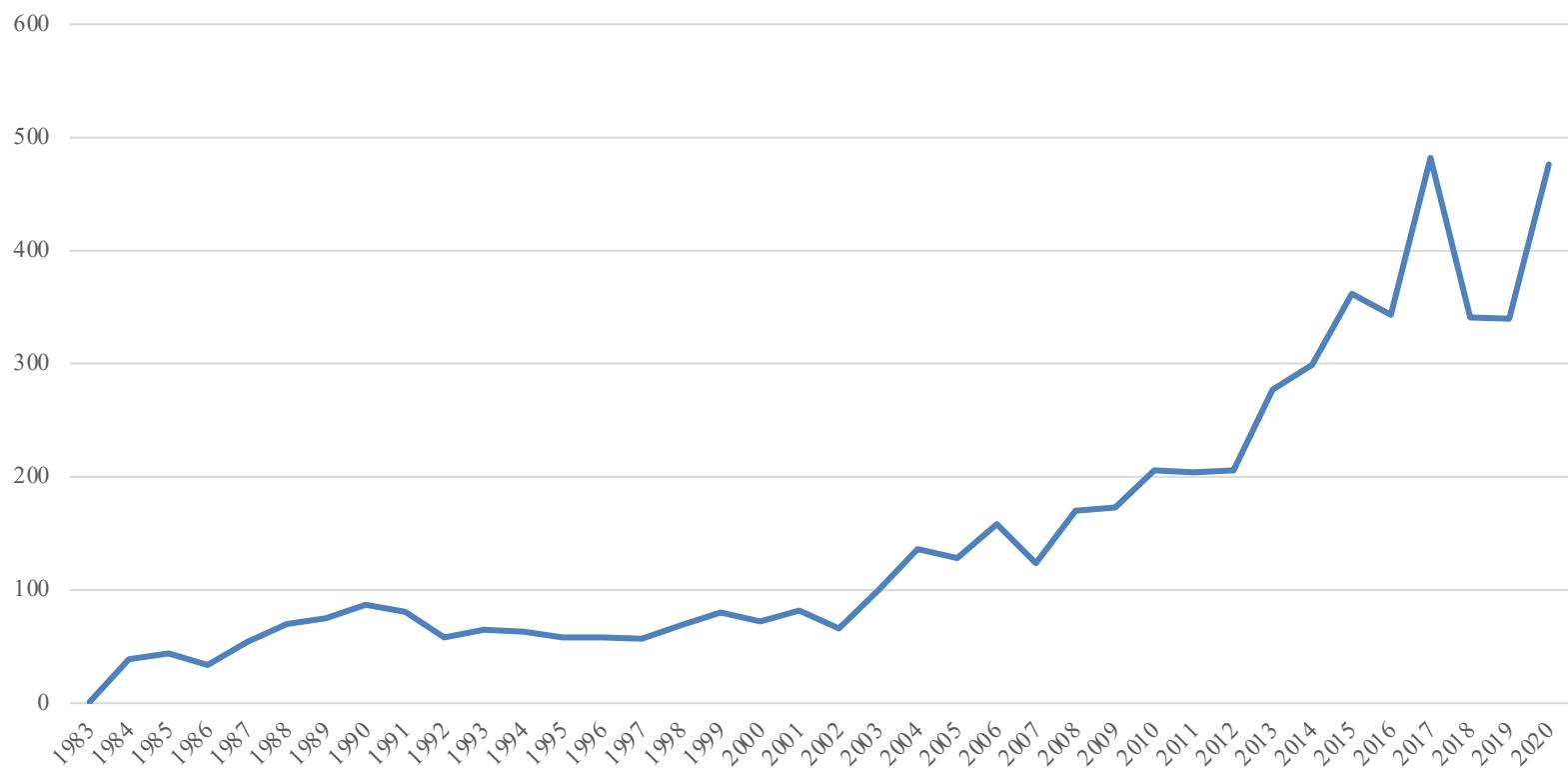
A	B	C	D	E	F
Generic Name	Trade Name	Date Designated	Orphan Designation	Orphan Designation Status	FDA Orphan Approval Status
(2S,3S,4E,6S,7R,10R)-7,10-Dihydroxy-3,7-dimethyl-12-oxo-2-[(2E,4E,6R)-6-(pyridin-2-yl)hepta-2,4-dien-2-yl]oxacyclododec-4-en-6-yl 4-methylpiperazine-1-carboxylate		6/22/17	Treatment of chronic myelomonocytic leukemia (CMML)	Designated	Not FDA Approved for Orphan Indication
(2S,4R)-1-(2-(3-acetyl-2-(methyl(pyrimidin-5-yl)-1H-indazol-1-yl)acetyl)-N-(6-bromopyridin-2-yl)-4-fluoropyrrolidine-2-carboxamide		11/2/17	Treatment of paroxysmal nocturnal hemoglobinuria	Designated	Not FDA Approved for Orphan Indication
(2-chloro-4-phenoxyphenyl)(4-(((3R,6S)-6-(hydroxymethyl)tetrahydro-2H-pyran-3-yl)amino)-7H-pyrrolo[2,3-d]pyrimidin-5-yl)methanone		12/8/20	Treatment of chronic lymphocytic leukemia (CLL) or small lymphocytic lymphoma (SLL)	Designated	Not FDA Approved for Orphan Indication
1-(2-Nitro-imidazolyl)-3-[18F]fluoro-2-propanol;1H-1-(3-[18F]fluoro-2-hydroxypropyl)-2-nitroimidazole		1/6/16	As a diagnostic for clinical management of soft tissue sarcoma, including rhabdomyosarcoma.	Designated	Not FDA Approved for Orphan Indication
2-(2-((2-(1H-benzimidazol-2-yl)ethyl)amino)ethyl)-N-[(3-fluoropyridin-2-yl)methyl]-1,3-oxazole-4-carboxamide trihydrochloride		6/17/19	Treatment of beta-thalassemia	Designated	Not FDA Approved for Orphan Indication
37,6?,7,7,12?-tetrahydroxy-5?-cholan-24-oic acid		10/22/20	treatment of progressive familial intrahepatic cholestasis	Designated	Not FDA Approved for Orphan Indication
4-(4-((6-(4-Chlorophenyl)-spiro[3.5]non-6-en-7-yl)methyl)-piperazin-1-yl)-N-((3-nitro-4-((2S)-1,4-dioxan-2-ylmethyl)amino)phenyl)sulfonyl)-2-(1H-pyrrolo[2,3-b]pyridin-5-yloxy)benzamide		8/28/20	Treatment of Chronic Lymphocytic Leukemia	Designated	Not FDA Approved for Orphan Indication
a non-replicating, recombinant adeno-associated virus serotype 9 (AAV9) containing the human Methyl Cpg Binding Protein 2B (MECP2B) cDNA under the control of a segment of the murine MECP2 promoter.		4/17/19	Treatment of Rett Syndrome (RTT)	Designated	Not FDA Approved for Orphan Indication
A recombinant plasmid-derived DNA-based vaccine targeting the E6 and E7 antigens of both HPV6 and HPV11		7/28/20	Treatment of recurrent respiratory papillomatosis	Designated	Not FDA Approved for Orphan Indication
a self-complementary adeno-associated virus serotype 9 (AAV9) vector expressing codon-optimized human SURF1 DNA coding sequence		10/20/20	Treatment of Leigh Syndrome.	Designated	Not FDA Approved for Orphan Indication
allogeneic CRISPR/Cas9-mediated genetically modified CAR T cells targeting CD19 antigen		2/16/21	Treatment of acute lymphoblastic leukemia	Designated	Not FDA Approved for Orphan Indication
antibody drug conjugate comprised of an anti-B cell maturation (BCMA) IgG1 humanized antibody conjugated covalently to the dibenzocyclooctyne (DBCO) noncleavable linker maytansinoid warhead		2/23/21	Treatment of multiple myeloma	Designated	Not FDA Approved for Orphan Indication
apadmatase alfa		9/28/20	Treatment of Sickle Cell Disease	Designated	Not FDA Approved for Orphan Indication
ascorbic acid		5/11/09	Treatment of Charcot-Marie-Tooth disease type 1A.	Designated	Not FDA Approved for Orphan Indication

Data Set (2/2)

- Total: 5849 records (treatment-disease)
- Generic Name (Column A)
 - Chemicals (substance and compound)
 - Medication, gene therapy or other therapies
 - Processed with mychem, pubchem and Google search
- Orphan Designation (Column D)
 - Processed with Medical NLP algorithms

Data Trend

Disease-Treatment Summary Trend (1983-2020)



Process Flow

A	B	C	D	E	F
Gene Name	Trait Name	Allele Designation	Protein Designation Name	Protein Designation Ref ID	Ref ID Protein Designation
CDLA	PBLN:CDL-00001	CDL-A-1A	Alpha 1A	PA10001	Ref ID Alpha 1A
CDLA	PBLN:CDL-00002	CDL-A-1B	Alpha 1B	PA10002	Ref ID Alpha 1B
CDLB	PBLN:CDL-00003	CDL-B-1A	Alpha 1A	PA10003	Ref ID Alpha 1A
CDLB	PBLN:CDL-00004	CDL-B-1B	Alpha 1B	PA10004	Ref ID Alpha 1B
CDLC	PBLN:CDL-00005	CDL-C-1A	Alpha 1A	PA10005	Ref ID Alpha 1A
CDLC	PBLN:CDL-00006	CDL-C-1B	Alpha 1B	PA10006	Ref ID Alpha 1B
CDLD	PBLN:CDL-00007	CDL-D-1A	Alpha 1A	PA10007	Ref ID Alpha 1A
CDLD	PBLN:CDL-00008	CDL-D-1B	Alpha 1B	PA10008	Ref ID Alpha 1B
CDLE	PBLN:CDL-00009	CDL-E-1A	Alpha 1A	PA10009	Ref ID Alpha 1A
CDLE	PBLN:CDL-00010	CDL-E-1B	Alpha 1B	PA10010	Ref ID Alpha 1B
CDLF	PBLN:CDL-00011	CDL-F-1A	Alpha 1A	PA10011	Ref ID Alpha 1A
CDLF	PBLN:CDL-00012	CDL-F-1B	Alpha 1B	PA10012	Ref ID Alpha 1B
CDLG	PBLN:CDL-00013	CDL-G-1A	Alpha 1A	PA10013	Ref ID Alpha 1A
CDLG	PBLN:CDL-00014	CDL-G-1B	Alpha 1B	PA10014	Ref ID Alpha 1B
CDLN	PBLN:CDL-00015	CDL-N-1A	Alpha 1A	PA10015	Ref ID Alpha 1A
CDLN	PBLN:CDL-00016	CDL-N-1B	Alpha 1B	PA10016	Ref ID Alpha 1B
CDLQ	PBLN:CDL-00017	CDL-Q-1A	Alpha 1A	PA10017	Ref ID Alpha 1A
CDLQ	PBLN:CDL-00018	CDL-Q-1B	Alpha 1B	PA10018	Ref ID Alpha 1B
CDLR	PBLN:CDL-00019	CDL-R-1A	Alpha 1A	PA10019	Ref ID Alpha 1A
CDLR	PBLN:CDL-00020	CDL-R-1B	Alpha 1B	PA10020	Ref ID Alpha 1B

Original Excel File

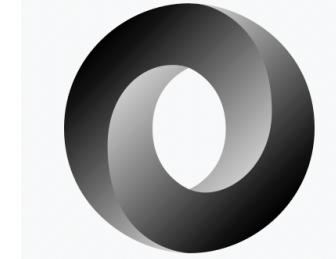


Manual
Data
curation

Multiple Myeloma
Treatment of multiple myeloma
Treatment of relapsed multiple myeloma
Treatment of acute myeloid leukemia
Treatment of myelodysplastic syndrome
Treatment of lymphoma
Treatment of hairy-cell leukaemia
Treatment of multiple myeloma
Prophylaxis for patients following discontinuation or suspected exposure to ibandronate

Neoplastic Process
Neoplastic Process

Processed Excel
File



JSON objects



Data Clean and Normalization (1/3) (column D)

- Example 1:
 - Treatment of paroxysmal nocturnal hemoglobinuria
paroxysmal nocturnal hemoglobinuria (C0024790;
Neoplastic Process)
- Example 2:
 - For immune reconstitution and prevention of graft versus host disease following allogeneic hematopoietic stem cell transplantation.
graft versus host disease (C0018133;Disease or Syndrome)

Data Clean and Normalization (2/3) (column D)

- Example 3:
 - Intravenous treatment of patients presenting with moderate to severe acetaminophen overdose
[Acetaminophen overdose \(C0572025; injury or poisoning\)](#)
- Example 4:
 - For use in post-exposure prophylaxis for occupational exposure to human immunodeficiency virus
[Exposure to Human immunodeficiency virus \(C1096519; injury or poisoning\)](#)

Data Clean and Normalization (3/3) (column D)

- Processing Types:
 - Disease or Syndrome
 - Neoplastic Process
 - Pathologic Function
 - Injury or Poisoning
 - Gene or Genome
 - Congenital Abnormality
 - Finding
 - Sign or Symptom
 - Intellectual Product
 - Anatomical Abnormality
 - Bacterium
 - Virus
 - Body Space or Junction
 - Mental or Behavioral Dysfunction
 - Cell Component
 - Eukaryote

Result – Column A

- Group 1 (51.92%; 3,037 records): PubChem, MyChem and Google search
 - Key: InChI Key or Substance ID
- Group 2 (48.08%; 2812 records): Not directly related to chemicals or did not find any match

Result – Column D

- Group 1 (96.6%; 5,650 records): Processed with type support

Record Id	Original text	Parsed Text	Text Type
1	Treatment of chronic myelomonocytic leukemia (CMML)	Chronic Myelomonocytic Leukemia	Neoplastic Process
2	Treatment of paroxysmal nocturnal hemoglobinuria	Paroxysmal nocturnal hemoglobinuria	Disease or Syndrome
3	Treatment of chronic lymphocytic leukemia (CLL) or small lymphocytic lymphoma (SLL)	Chronic Lymphocytic Leukemia	Neoplastic Process
3	Treatment of chronic lymphocytic leukemia (CLL) or small lymphocytic lymphoma (SLL)	Small Lymphocytic Lymphoma	Neoplastic Process
3	Treatment of chronic lymphocytic leukemia (CLL) or small lymphocytic lymphoma (SLL)	SLL	Gene or Genome
4	As a diagnostic for clinical management of soft tissue sarcoma, including rhabdomyosarcoma.	Rhabdomyosarcoma	Neoplastic Process
4	As a diagnostic for clinical management of soft tissue sarcoma, including rhabdomyosarcoma.	Soft tissue sarcoma	Neoplastic Process
4	As a diagnostic for clinical management of soft tissue sarcoma, including rhabdomyosarcoma.	Diagnostic	Diagnostic Procedure
5	Treatment of beta-thalassemia	BETA-THALASSEMIA	Disease or Syndrome

- Group 2 (3.4%; 199 records): Not directly related to disease or did not find any match

2247 For promotion of cutaneous wound healing in extreme burn treatment protocols.

2321 Prevention of graft rejection following solid organ transplantation

2611 Mobilization of hematopoietic stem cells to the peripheral blood for collection and subsequent transplant

3006 Prevention of graft rejection following solid organ transplantation

3090 For the promotion of early engraftment in bone marrow transplantation.

For promotion of cutaneous wound healing in extreme burn treatment protocols.

Prevention of graft rejection following solid organ transplantation

Mobilization of hematopoietic stem cells to the peripheral blood for collection and subsequent transplant

Prevention of graft rejection following solid organ transplantation

For the promotion of early engraftment in bone marrow transplantation.

Discussion

- JSON format

Search Orphan Drug Designations and Approvals

[FDA Home](#) [Developing Products for Rare Diseases & Conditions](#)

Generic Name:	(2S,3S,4E,6S,7R,10R)-7,10-Dihydroxy-3,7-dimethyl-12-oxo-2-[(2E,4E,6R)-6-(pyridin-2-yl)hepta-2,4-dien-2-yl]oxacyclododec-4-en-6-yl 4-methylpiperazine-1-carboxylate
Date Designated:	06/22/2017
Orphan Designation:	Treatment of chronic myelomonocytic leukemia (CMML)
Orphan Designation Status:	Designated
FDA Orphan Approval Status:	Not FDA Approved for Orphan Indication
Sponsor:	H3 Biomedicine Inc. (US Research Subsidiary of Eisai Co., Ltd.) 300 Technology Square Floor 5 Cambridge, Massachusetts 02139 United States

The sponsor address listed is the last reported by the sponsor to OOPD.

Discussion

- Index rule

Modify indexing rules

Field: gene

Path: annotations.gene

Index this field

Search this field by default

Change type text ▾

```
{  
  "type": "text",  
  "copy_to": [  
    "all"  
  ]  
}
```

- ▶ Enable index allows a field to be searchable. If indexing is disabled, values are still stored and returned in results, but they can't be directly queried. Indexing takes disk space and can also impact performances, only index fields which make sense to query.
- ▶ When **Search by default** is enabled, field can be searched without specifying the full path. Note: `_id` field is an exception, path is not required.
Ex:
 - When searching by default is disabled, searching `gene` field requires to specify the full path:
`/query?q=annotations.gene:value_to_search`
 - When searching by default is enabled, path can be omitted, `value_to_search` will be searched in all fields declared as searchable by default.
`/query?q=value_to_search`.
- ▶ ElasticSearch field data type can be changed if needed. See [Field datatypes](#) for more information. Note: only core datatypes are available in this list
- ▶ Field definition can also be manually specified in the text box, using JSON notation, for more advanced usage.

 Cancel

 OK



Thank you