

OHDSI Oncology Workgroup - cancer drug name research

25 July 2017

Don O'Hara

Questions:

1. What oncology drugs are being used "in the field", and how are they being reported?
2. How well represented are they in the standard drug vocabularies?

Methods:

1. Get list of cancer drugs from NCI:

<https://www.cancer.gov/about-cancer/treatment/drugs>
"A to Z List of Cancer Drugs"

- Downloaded list of 497 drug names
- Parsed into 1,003 tokens

2. Compared to recent (May 2017) OHDSI Concept table

- Extracted 956,041 drugnames, for restricted set of vocabulary_ids
 - domain_id = 'Drug'
 - vocabulary_id in ('CPT4', 'HCPCS', 'NDC', RxNorm')
 - Count of concepts having domain_id = 'Drug' = 3,001,184)
- Parsed into 6,962,350 tokens

domain_id	vocabular_id	count	selected
Drug	AMT	72,216	
Drug	ATC	6,037	
Drug	BDPM	30,823	
Drug	CIEL	8,043	
Drug	Cohort	12	
Drug	CPT4	105	X
Drug	dm+d	149,219	
Drug	DPD	167,438	
Drug	EphMRA ATC	895	
Drug	GCN_SEQNO	29,659	
Drug	Gemsript	224,570	

Drug	GRR	140,664	
Drug	HCPCS	1,076	X
Drug	MeSH	4,635	
Drug	Multum	9,770	
Drug	NDC	679,882	X
Drug	NDFRT	37,008	
Drug	NFC	692	
Drug	OXMIS	3	
Drug	Read	63	
Drug	RxNorm	274,978	X
Drug	RxNorm Extension	617,965	
Drug	SNOMED	323,475	
Drug	SPL	167,259	
Drug	VA Class	486	
Drug	VA Product	54,211	

3. Compared NCI tokens vs OHDSI drug tokens, excluding noise words:

acetate
acid
aerosol
and
bromide
calcium
citrate
disodium
for
human
hydrochloride
implant
injection
iodine
mesylate
monohydrate
phosphate
powder
recombinant

sodium
sterile
sulfate
tablets
tartrate
top
vaccine

Results:

Metric	N	Pct
NCI drug names	497	
Matched to some OHDSI Drug concept	428	86.1
Not matched to any OHDSI Drug concept	69	13.9

4. No match on OHDSI for these NCI drug named (n=69) :

5-fu (fluorouracil--topical)
alunbrig (brigatinib)
avelumab
bavencio (avelumab)
beacopp
bep
brigatinib
bumel
cabometyx (cabozantinib-s-malate)
cabozantinib-s-malate
capox
carac (fluorouracil--topical)
carboplatin-taxol
cem
cev

chlorambucil-prednisone
chop
cmf
cobimetinib
cometriq (cabozantinib-s-malate)
copdac
copp
copp-abv
cotellic (cobimetinib)
cryo-em
cvp
durvalumab
efudex (fluorouracil--topical)
fec
fluoroplex (fluorouracil--topical)
fluorouracil--topical
foia
folfiri
folfiri-bevacizumab
folfiri-cetuximab
folfirinox
folfox
fu-lv
gemcitabine-cisplatin
gemcitabine-oxaliplatin
hyper-cvad
imfinzi (durvalumab)
imlygic (talimogene laherparepvec)
jeb
kisqali (ribociclib)
lartruvo (olaratumab)
midostaurin
mopp
oepa
olaratumab
oppa
pcv
r-chop

r-cvp
r-epoch
r-ice
randomization
ribociclib
rss
rydapt (midostaurin)
tac
talimogene laherparepvec
tolak (fluorouracil--topical)
top
tpf
vamp
veip
xeliri
xelox

Some false negatives due to parsing

NCI Drug name	token
5-fu (fluorouracil injection)	5-fu
5-fu (fluorouracil injection)	fluorouracil
5-fu (fluorouracil injection)	injection
5-fu (fluorouracil--topical)	5-fu
5-fu (fluorouracil--topical)	fluorouracil--topical

NCI Drug name	Matches
5-fu (fluorouracil injection)	108 matched tokens
5-fu (fluorouracil--topical)	0 matches

nci drug_name	token	OHDSI concept_name	concept_code	vocabulary_id
5-fu (fluorouracil injection)	fluorouracil	fluorouracil 50mg/g TOPICAL CREAM	665300249	NDC
5-fu (fluorouracil injection)	fluorouracil	fluorouracil 50mg/g TOPICAL CREAM	516724118	NDC
5-fu (fluorouracil injection)	fluorouracil	fluorouracil 50mg/mL INTRAVENOUS INJECTION, SOLUTION	167290276	NDC
5-fu (fluorouracil injection)	fluorouracil	Injection, fluorouracil, 500 mg	J9190	HCPCS
5-fu (fluorouracil injection)	fluorouracil	fluorouracil 50mg/mL INTRAVENOUS INJECTION, SOLUTION	633230117	NDC
abiraterone acetate	abiraterone	abiraterone Pill	1163657	RxNorm
abiraterone acetate	abiraterone	abiraterone Oral Tablet	1100074	RxNorm
abiraterone acetate	abiraterone	abiraterone acetate 250 MG	1100073	RxNorm
abiraterone acetate	abiraterone	abiraterone acetate 250 MG [Zytiga]	1100077	RxNorm
abiraterone acetate	abiraterone	abiraterone acetate	1100071	RxNorm
abiraterone acetate	abiraterone	methotrexate 10mg/.4mL SUBCUTANEOUS INJECTION, SOLUTION [otrexup]	544360010	NDC
abiraterone acetate	abiraterone	methotrexate 15 mg/1 TABLET, FILM COATED [Trexall]	51285369	NDC
abiraterone acetate	abiraterone	methotrexate sodium 25mg/mL INTRA-ARTERIAL; INTRAMUSCULAR; INTRATHECAL; INTRAVENOUS INJECTION, SOLUTION [methotrexate]	7033673	NDC
abiraterone acetate	abiraterone	methotrexate 2.5mg/1 ORAL TABLET	422540110	NDC
abiraterone acetate	abiraterone	methotrexate 25mg/.4mL SUBCUTANEOUS INJECTION, SOLUTION [otrexup]	544360025	NDC
abiraterone acetate (paclitaxel albumin-stabilized nanoparticle formulation)	formulation	Influenza virus vaccine, pandemic formulation, live, for intranasal use	90664	CPT4
abiraterone acetate (paclitaxel albumin-stabilized nanoparticle formulation)	formulation	Injection, daunorubicin citrate, liposomal formulation, 10 mg	J9151	HCPCS
abiraterone acetate (paclitaxel albumin-stabilized nanoparticle formulation)	formulation	Influenza virus vaccine, pandemic formulation, split virus, preservative free, for intramuscular use	90666	CPT4
abiraterone acetate (paclitaxel albumin-stabilized nanoparticle formulation)	formulation	Influenza virus vaccine, pandemic formulation, split virus, for intramuscular use	90668	CPT4
abiraterone acetate (paclitaxel albumin-stabilized nanoparticle formulation)	formulation	Influenza virus vaccine, pandemic formulation, split virus, adjuvanted, for intramuscular use	90667	CPT4

5. Sample matched tokens:

Next Steps:

1. Expand list of "currently used oncology drugs"
2. Refine token matching
3. Eliminate noise words ("formulation")

4. Review what vocabulary_ids should be included in the "OHDSI drug" set
5. Determine if there are other vocabularies not in the current OHDSI vocabulary set that may contain oncology treatment concepts.
6. Refresh local copy of OHDSI concepts