

# **WEB-RADR WP2A REPORT**

## **Report on Data Collection**

Type of document:	Report
Version:	0.1
Date:	17 November 2016
Authors:	Carrie Pierce

## WP2A Data Collection Report

### Introduction

WP2A has been tasked with developing a collaborative workspace for visualizing and analyzing social media data pertaining to WEB-RADR products. To meet this objective, Epidemico's MedWatcher Social platform has been adapted to monitor WEB-RADR products in three European languages – English, French, and Spanish.

MedWatcher Social is a social media monitoring platform that collects publicly available data from Internet social media sources. MedWatcher Social is designed to complement existing safety efforts by supporting post-marketing safety surveillance. Filtering and classification algorithms are applied to identify content relating to adverse event experiences with medical products (namely drugs, devices, and vaccines). The filtered, de-identified, and aggregated data are then made available to end users for review and analysis via an interactive Web-based visualization dashboard. This document provides an overview of the data that has been collected via the MedWatcher Social system to date.

### Data Collection Statistics

#### *Data Acquisition*

As of October 2016, we have collected 2,823,447 posts mentioning any of the 113 medical products selected by the WEB-RADR consortium at the beginning of the project (Appendix A). These posts have resulted from both prospective data collection since September 2014 as well as a historical data acquisition (January 1, 2012-December 31, 2014) intended to facilitate analysis and evaluation activities in WP2B and WP4. The majority of posts (83%) were acquired from Twitter.

#### *Automated Classification*

Over half (54.10%) of all posts collected for the project were automatically classified as spam or irrelevant posts, leaving 1567241 “meaningful mentions” – posts that legitimately refer to a medical product and do not contain “bot-like” language. Of these mentions, 30,300 (1.94%) were classified as posts with resemblance to adverse events (Proto-AEs).

### New Data Sources

In light of the loss of access to post-level Facebook data in October 2015, it has become important for the consortium to consider additional data sources to include in data collection and analyses in order to ensure the applicability and generalizability of WEB-RADR results beyond the culmination of the project. The following data sources have been added to the data acquisition process with the intent of determining their value as possible sources of patient-generated drug safety information. WP2A will work with WP2B and WP4 to determine a plan for delivery of additional datasets and analysis of this additional data in Year 3 of the project.

### *Reddit*

Reddit is an open-source online community that was launched in June 2005. It consists of 9,330 active themed forums, called subreddits. Reddit users (“redditors”) can post stories within the subreddits or comment on stories. Reddit has 202 million monthly unique visitors from 208 different countries.

As of October 2016, historical Reddit posts spanning the first two years of the WEB-RADR project have been acquired from a publicly available data corpus published by a Reddit user. Prospective data for the remaining year of the project will be acquired in real-time using Reddit’s Application Programming Interface (API).

To date, the Reddit consists of 157,215 posts containing reference to any of the WEB-RADR products. From this corpus, 18% (n = 28,597) of all posts were tagged “junk” – posts containing irrelevant language – leaving 128,618 “meaningful mentions.” Of these mentions, 20,881 posts (16% of mentions and 13% of all posts collected) were tagged as Proto-AEs.

Geographic information at the country level was only available for 0.6% of all reddit posts. Of these, all authors were from the United States or Canada.

### *Forums*

Patient forums are digital message boards or online communities that contain conversations (i.e., threads) dedicated to specific therapeutic areas or health concerns. While these data sources typically yield lower volume than social networks, it is suspected that the data from patient forums may contain more insights due to the typical length of these posts. Furthermore, forums provide a dedicated space for patients to discuss medical issues with other patients who are living with similar conditions or taking similar treatments.

Data from patient forums are available via authorized third-party data vendors or in-house aggregation machines while adhering to each site’s terms of use. As of October 2016, we have collected 199,790 posts from various patient forums. Of these posts, 21% (n = 41,853) posts were tagged as “junk”, leaving 157,937 “meaningful mentions.” Of these mentions, 54,429 posts (34% of mentions and 27% of all posts collected) were tagged as Proto-AEs.

The following forums have each generated more than 20 Proto-AEs thus far, and therefore may be considered “higher value” data sources than others.

<b>Forum Domain</b>	<b>Number of Proto-AEs</b>
breastcancer.org	1188
babycenter.com	232
her2support.org	162
fertilityfriends.co.uk	125
cancer.org	116
healthunlocked.com	98
drugs-forum.com	97

healingwell.com	185
medschat.com	91
dailystrength.org	86
medhelp.org	78
crazymeds.us	63
breastcancercare.org.uk	57
patient.co.uk	49
myelomabeacon.com	48
colonclub.com	45
tudiabetes.org	42
cancercompass.com	39
babiesbase.com	32
fertilethoughts.com	32
proboards.com	31
cancerforums.net	29
hystersisters.com	29
talkingcity.com	27
longecity.org	26
diabetessupport.co.uk	25
mdjunction.com	25
medhelp.org	23
navigatingcancer.com	23
whattoexpect.com	22
childrenwithdiabetes.com	21
thebump.com	21
macmillan.org.uk	20

### *Inspire*

Inspire: Health and Wellness Support Groups and Communities is an online patient community with dedicated groups for over 190 distinct therapeutic areas. These disease-specific communities are developed in partnership with patient advocacy groups and user-generated content from these communities are in turn leveraged for secondary research by life science organizations. Inspire has over 700,000 monthly unique visitors.

WP2A leadership has worked with Inspire directly to acquire a corpus of data throughout the course of the project, as a potential replacement for Facebook data. Inspire will provide historical data spanning the first two years of the project as a one-time deliverable; prospective data for the remaining year of the project will be provided on a quarterly basis.

As of October 2016, the first batch of Inspire data has been acquired and processed. This corpus consists of 45,833 posts containing reference to any of the WEB-RADR products. From this corpus, 13% (n = 6,002) posts were tagged “junk”, leaving 39,831

“meaningful mentions.” Of these mentions, 5,489 posts (14% of mentions and 12% of all posts collected) were tagged as Proto-AEs.

Each Inspire post is accompanied by metadata that includes the author’s country code. Out of the total data corpus, 5.87% of posts were published by Inspire community members in the European Union, while 6.43% of all Proto-AEs were published by EU inhabitants. The table displays the total volume of posts within each EU member state.

EU REPRESENTATION		
Country	Total Posts	Number of Proto-AEs
Austria	24	1
Belgium	10	1
Bulgaria	3	0
Croatia	5	0
Cyprus	5	0
Czech Republic	1	0
Denmark	38	5
Estonia	7	0
Finland	4	0
France	28	5
Germany	55	8
Greece	16	2
Hungary	3	0
Ireland	218	28
Italy	51	2
Latvia	1	0
Lithuania	4	0
Luxemburg	5	0
Malta	9	0
Netherlands	84	7
Poland	35	4
Portugal	23	0
Romania	8	2
Slovakia	16	1
Slovenia	21	4
Spain	111	7
Sweden	20	1
United Kingdom	1886	275
<b>Total</b>	<b>2691</b>	<b>353</b>

## Appendix A

### WEB-RADR Product Search Terms

Product Name	Active Ingredients / Generics	Search Terms
acarbose	acarbose	acarbose
Aclasta	zoledronic acid	Aclasta
Adalat	nifedipine	Adalat
Adefin	nifedipine	Adefin
Afinitor	everolimus	Afinitor
aldesleukin	aldesleukin	aldesleukin
alemtuzumab	alemtuzumab	alemtuzumab
aliskiren	aliskiren	aliskiren
Allurene	dienogest	Allurene
Alpharadin	radium Ra 223 dichloride	Alpharadin
amlodipine and valsartan	amlodipine,valsartan	amlodipine and valsartan
amlodipine-benazepril	amlodipine,benazepril	amlodipine and benazepril
amlodipine, valsartan and hydrochlorothiazide	amlodipine,valsartan,hydrochlorothiazide	amlodipine valsartan hydrochlorothiazide
Anafranil	clomipramine	Anafranil
Anastrozol	anastrozole	Anastrozol
anastrozole	anastrozole	anastrozole
Apidra	insulin glulisine	Apidra
Aranesp	darbepoetin alfa	Aranesp
Arcapta/Onbrez	Indacaterol Maleate	Onbrez, Arcapta
Aredia	pamidronate	Aredia
Arimidex	anastrozole	Arimidex
artemether and lumefantrine	artemether,lumefantrine	lumefantrine
AscoTop	zolmitriptan	AscoTop
Atenolol	atenolol	Atenolol
Aubagio	teriflunomide	Aubagio
Avlocardyl	propranolol	Avlocardyl
Axanum	esomeprazole,acetylsalicylic acid	Axanum

Baclofen	Baclofen	Baclofen
basiliximab	basiliximab	basiliximab
Betaferon	interferon beta-1b	Betaferon
Betaseron	interferon beta-1b	Betaseron
Betazide	metoprolol tartrate	Betazide
bicalutamide	bicalutamide	bicalutamide
Brilinta	Ticagrelor	Brilinta
Brilique	Ticagrelor	Brilique
budesonide	budesonide	budesonide
Budicort	budesonide	Budicort
Cabazitaxel	Cabazitaxel	Cabazitaxel
canakinumab	canakinumab	canakinumab
Caprelsa	vandetanib	Caprelsa
carbamazepine	carbamazepine	carbamazepine
Cardioxane	dexrazoxane	Cardioxane
Casodex	bicalutamide	Casodex
Ceftaroline foramil	Ceftaroline foramil	Ceftaroline
ceritinib	ceritinib	ceritinib
Certican	everolimus	Certican
Chronadalat	nifedipine	Chronadalat
Chronadalate	nifedipine	Chronadalate
cinacalcet	cinacalcet	cinacalcet
Claritrast	iopromide	Claritrast
Clarograf	iopromide	Clarograf
Climagest	norethisterone,estradiol valerate	Climagest
Climaval	estradiol valerate	Climaval
Climesse	norethisterone,estradiol valerate	Climesse
clomipramine	clomipramine	clomipramine
clopidogrel	clopidogrel	clopidogrel
clozapine	clozapine	clozapine,clozpine

Clozaril	clozapine	Clozaril
Co-Diovan	valsartan,hydrochlorothiazide	Co Diovan, Cotareg
Compesk	interferon beta-1b	Compesk
Comtan	entacapone	Comtan
Cosentyx	secukinumab	Cosentyx
Cosudex	bicalutamide	Cosudex
cyclosporin	cyclosporin	cyclosporin
darbepoetin alfa	darbepoetin alfa	darbepoetin
deferasirox	deferasirox	deferasirox
deferoxamine	deferoxamine	deferoxamine
Denosumab	Denosumab	Denosumab
Desferal	deferoxamine	Desferal, Desferin
dexrazoxane	dexrazoxane	dexrazoxane
Diclofenac	diclofenac	diclofenac
dienogest	dienogest	dienogest
Diovan	valsartan	Diovan, Tareg
Diprivan	propofol	Diprivan
Dociton	propranolol	Dociton
dronedarone	dronedarone	dronedarone
entacapone	entacapone	entacapone
Entresto	valsartan,sacubitril	Entresto
esomeprazole and acetylsalicylic acid	esomeprazole,acetylsalicylic acid	esomeprazole and acetylsalicylic acid
Estraderm	estradiol hemihydrate	Estraderm
Estradiol	estradiol	Estradiol
estradiol hemihydrate	estradiol hemihydrate	estradiol hemihydrate
estradiol valerate	estradiol valerate	estradiol valerate
ethinylestradiol and gestodene	ethinylestradiol,gestodene	ethinylestradiol and gestodene
Eucreas	vildagliptin,metformin	Eucreas
everolimus	everolimus	everolimus
Exelon	rivastigmine	Exelon



Exelon Patch	rivastigmine transdermal	exelon patch
Exforge	amlodipine,valsartan	Exforge
Exforge HCT	amlodipine,valsartan,hydrochlorothiazide	Exforge HCT
Exjade	deferasirox	Exjade
Extavia	interferon beta-1b	Extavia
famciclovir	famciclovir	famciclovir
Famvir	famciclovir	Famvir, Oravir
Fanapt	iloperidone	Fanapt
Farydak	panobinostat	Farydak
Fedra	ethinylestradiol,gestodene	Fedra
Feloday	felodipine	Feloday
Felodipin	felodipine	Felodipin
felodipine	felodipine	felodipine
Felodur	felodipine	Felodur
Femara	letrozole	Femara
Femiane	ethinylestradiol,gestodene	Femiane
Femodette	ethinylestradiol,gestodene	Femodette
Feronia	interferon beta-1b	Feronia
Ferro Gyn	ferrous glycine sulfate,folic acid	Ferro Gyn
Ferro Sanol	ferrous glycine sulfate	Ferro Sanol
Ferro Sanol Gyn	ferrous glycine sulfate,folic acid	Ferro Sanol Gyn
ferrous glycine sulfate	ferrous glycine sulfate	ferrous glycine
ferrous glycine sulfate and folic acid	ferrous glycine sulfate,folic acid	ferrous glycine sulfate and folic acid
filgrastim	filgrastim	filgrastim
fingolimod	fingolimod	fingolimod
Fluenz	influenza	Fluenz
FluMist	influenza	FluMist,flu spray
fluvastatin	fluvastatin	fluvastatin
Galvus	vildagliptin	Galvus
Gilenya	fingolimod	Gilenya

Gitsalat	nifedipine	Gitsalat
Glicobase	acarbose	Glicobase
Glivec	imatinib	Glivec
Glucobay	acarbose	Glucobay
Gluconase	acarbose	Gluconase
Glucor	acarbose	Glucor
Glumida	acarbose	Glumida
glycopyrronium bromide	glycopyrronium bromide	glycopyrronium
Gynera	ethinylestradiol,gestodene	Gynera
Gynovin	ethinylestradiol,gestodene	Gynovin, Gynoden
Ilaris	canakinumab	Ilaris
iloperidone	iloperidone	iloperidone
imatinib	imatinib	imatinib
Imodium Multi-Symptom	loperamide and simethicone	Imodium
indacaterol and glycopyrronium	Indacaterol Maleate, glycopyrronium bromide	indacaterol and glycopyrronium
Indacaterol Maleate	Indacaterol Maleate	Indacaterol
Inderal	propranolol	Inderal
Inderalici	propranolol	Inderalici
insulin glargine	insulin glargine	glargine
insulin glulisin	insulin glulisine	glulisine
interferon beta-1b	interferon beta-1b	interferon beta 1b
Invega	paliperidone	Invega
Invega Sustenna	paliperidone palmitate	Invega Sustenna
iopromide	iopromide	iopromide
Istubal	tamoxifen	Istubal
Jakavi	Ruxolitinib Phosphate	Jakavi
Jevtana	Cabazitaxel	Jevtana
Keppra	levetiracetam	Keppra
Kombiglyze	saxagliptin	Kombiglyze
Lamisil	terbinafine	Lamisil

Lantus	insulin glargine	Lantus
Lemtrada	alemtuzumab	Lemtrada
Lescol	fluvastatin	Lescol
letrozole	letrozole	letrozole
Levetiracetam	levetiracetam	Levetiracetam
Levitra	vardenafil hydrochloride	Levitra
levodopa/entacapone/carbidopa	levodopa,entacapone,carbidopa	levodopa entacapone carbidopa
Lioresal	Baclofen	Lioresal
lixisenatide	lixisenatide	lixisenatide
Lodene	ethinylestradiol,gestodene	Lodene
Logest	ethinylestradiol,gestodene	Logest
loperamide and simethicone	loperamide and simethicone	loperamide and simethicone
Lotrel	amlodipine,benazepril	Lotrel
Lucentis	ranibizumab	Lucentis
lumiracoxib	lumiracoxib	lumiracoxib
Lunafem	ethinylestradiol,gestodene	Lunafem
Lyxumia	lixisenatide	Lyxumia
Meliane	ethinylestradiol,gestodene	Meliane
Meloden	ethinylestradiol,gestodene	Meloden
methylphenidate	methylphenidate	methylphenidate
metoprolol tartrate	metoprolol tartrate	metoprolol
Mimpara	cinacalcet	Mimpara
Minigeste	ethinylestradiol,gestodene	Minigeste
Multaq	dronedarone	Multaq
mycophenolic acid	mycophenolic acid	mycophenolic acid
Myfortic	mycophenolic acid	Myfortic
naproxen and esomeprazole	naproxen,esomeprazole	naproxen and esomeprazole
Navoban	tropisetron	Navoban
Neoral	cyclosporin	Neoral
Neulasta	pegfilgrastim	Neulasta

Neupogen	filgrastim	Neupogen
Nexavar	sorafenib	Nexavar
nifedipine	nifedipine	nifedipine
Niferex	ferrous glycine sulfate	Niferex
nilotinib	nilotinib	nilotinib
Nitroderm	nitroglycerine	Nitroderm, Nitriderm
nitroglycerine	nitroglycerine	nitroglycerine
Nolvadex	tamoxifen	Nolvadex
norethisterone and estradiol valerate	norethisterone,estradiol valerate	norethisterone and estradiol valerate
Nplate	romplostim	Nplate
Obsidan Fe++	ferrous glycine sulfate	Obsidan
octreotide	octreotide	octreotide
omalizumab	omalizumab	omalizumab
Onglyza	saxagliptin	Onglyza
oxcarbazepine	oxcarbazepine	oxcarbazepine
paliperidone	paliperidone	paliperidone
paliperidone palmitate	paliperidone palmitate	paliperidone palmitate
pamidronate	pamidronate	pamidronate
panitumumab	panitumumab	panitumumab
panobinostat	panobinostat	panobinostat
pasireotide	pasireotide	pasireotide
pegfilgrastim	pegfilgrastim	pegfilgrastim
Perfudal	felodipine	Perfudal
Pertensal	nifedipine	Pertensal
pilocarpine	pilocarpine	pilocarpine
Plavix	clopidogrel	Plavix
Plendil	felodipine	Plendil, Flodil
Ponesta	zolmitriptan	Ponesta
Prandase	acarbose	Prandase
Precose	acarbose	precose

Preslow	felodipine	Preslow
Prevex	felodipine	Prevex
Prexige	lumiracoxib	Prexige
Proleukin	aldesleukin	Proleukin
Prolia	Denosumab	Prolia
propofol	propofol	propofol
Propranolol	propranolol	propranolol
Proscope	iopromide	Proscope
Pulmaxan	budesonide	Pulmaxan
Pulmicort	budesonide	pulmicort
radium Ra 223 dichloride	radium Ra 223 dichloride	radium dichloride,Ra 223
ranibizumab	ranibizumab	ranibizumab
Rasilez	aliskiren	Rasilez
Regorafenib	Regorafenib	Regorafenib
Remidex	anastrozole	Remidex
Riamet	artemether,lumefantrine	Riamet
Ritalin	methylphenidate	Ritalin
rivastigmine	rivastigmine	rivastigmine
rivastigmine transdermal	rivastigmine transdermal	rivastigmine transdermal
romplostim	romplostim	romplostim
Ruxolitinib Phosphate	Ruxolitinib Phosphate	Ruxolitinib
sacubitril and valsartan	valsartan,sacubitril	sacubitril and valsartan
Salagen	pilocarpine	Salagen
Sandimmun	cyclosporin	Sandimmun
Sandostatin	octreotide	Sandostatin, Sandostatine
saxagliptin	saxagliptin	saxagliptin
Sebivo	telbivudine	Sebivo
secukinumab	secukinumab	secukinumab
Seebri	glycopyrronium bromide	Seebri
Seloken	metoprolol tartrate	Seloken, Beloken

Selozok	metoprolol tartrate	Selozok
Signifor	pasireotide	Signifor
Simulect	basiliximab	Simulect
sorafenib	sorafenib	sorafenib
Spartofer	ferrous glycine sulfate	Spartofer
Spesicor	metoprolol tartrate	Spesicor
Spirocort	budesonide	Spirocort
Stabilic	acarbose	Stabilic
Stalevo	levodopa,entacapone,carbidopa	Stalevo
Staxyn	vardeafil hydrochloride	Staxyn
Stilnox	zolpidem	Stilnox
Stivar	Regorafenib	Stivar
Stivarga	Regorafenib	Stivarga
Sumial	propranolol	Sumial
Tactevo	influenza	Tactevo
tamoxifen	tamoxifen	tamoxifen
Tasigna	nilotinib	Tasigna
Tegretol	carbamazepine	Tegretol
Tekturna	aliskiren	Tekturna
telbivudine	telbivudine	telbivudine
Tenormin	atenolol	Tenormin, Tenormine
terbinafine	terbinafine	terbinafine
teriflunomide	teriflunomide	teriflunomide
Ticagrelor	Ticagrelor	Ticagrelor
Tobi	tobramycin	Tobi
tobramycin	tobramycin	tobramycin
Topamax	topiramate	Topamax
topiramate	topiramate	topiramate
Trileptal	oxcarbazepine	trileptal
tropisetron	tropisetron	tropisetron

Ultibro	Indacaterol Maleate, glycopyrronium bromide	Ultibro
Ultravist	iopromide	Ultravist
Valodex	tamoxifen	Valodex
valsartan	valsartan	valsartan
valsartan and hydrochlorothiazide	valsartan, hydrochlorothiazide	valsartan and hydrochlorothiazide
vandetanib	vandetanib	vandetanib
vardenafil	vardenafil hydrochloride	vardenafil
Vectibix	panitumumab	Vectibix
verteporfin	verteporfin	verteporfin
vildagliptin	vildagliptin	vildagliptin
vildagliptin and metformin	vildagliptin, metformin	vildagliptin and metformin
Vimovo	naproxen, esomeprazole	Vimovo
Visabelle	dienogest	Visabelle
Visanna	dienogest	Visanna
Visanne	dienogest	Visanne
Visannette	dienogest	Visannette
Visudyne	verteporfin	Visudyne
Vivanza	vardenafil hydrochloride	Vivanza
Vivelle	estradiol	Vivelle, Vivelledot, Estradot
Voltarol	diclofenac	Voltarol
Votubia	everolimus	Votubia
Xgeva	Denosumab	Xgeva
Xofigo	radium Ra 223 dichloride	Xofigo
Xolair	omalizumab	Xolair
Yaila	vardenafil hydrochloride	Yaila
Zactima	vandetanib	Zactima
Zaltrap	ziv-aflibercept	Zaltrap
Zinforo	Ceftaroline foramil	Zinforo
ziv-aflibercept	ziv-aflibercept	Aflibercept
zoledronic acid	zoledronic acid	zoledronic acid

zolmitriptan	zolmitriptan	zolmitriptan
Zolpidem	zolpidem	Zolpidem
Zometa	zoledronic acid	Zometa
Zomig	zolmitriptan	Zomig
Zomigon	zolmitriptan	Zomigon
Zomigoro	zolmitriptan	Zomigoro
Zykadia	ceritinib	Zykadia