SWAP Algorithm Verification

Test outbreak library

1/7/2016
Los Alamos National Laboratory
Data Dreamers

Table of Contents

Objective

The SWAP web app uses an algorithm to assign percentage similarity/match scores for user input and a historical outbreak from the SWAP library. This algorithm requires disease specific properties and specific weights associated with each property. Outbreaks that are not included in the library can be used to test the properties and weights of a disease specific algorithm. We collected information for test outbreaks and used it as input for the SWAP to evaluate SWAP output. The expected result for the test was that the top 1-5 matching historical outbreaks would have case counts and durations that reflect the scenario in the test outbreak.

We also evaluated a lower limit or threshold match value for each disease for a relevant match. The relevant match cut off indicates the percentage match below which the similarities to historic outbreaks were found to be "not relevant". However, this does not mean all matches above a given percentage will be relevant to the user input scenario. For example, if the disease has a cut off threshold of 70% and the top 5 SWAP contains 84-64% match score. The results interpretation may be as follows: outbreaks with scores below 70% can be ignored. After evaluating historic outbreaks above 70%, it may be determined only top 1 and top 2 are reasonable matches. The third match despite a match score of >70% may not be a relevant match due to improbably high case number and/or duration. The disease specific tables contain our reasoning for non-matching scenarios. The notes session of the tables are free text.

This document contains user input information for test outbreaks used for all of the SWAP diseases, the similarity/match score range for the top 5 matches (in percentage), the number of expert determined acceptable matches in top 5 and the reasoning (notes). A web link to the test outbreak data source is also provided.

Examples of test outbreak data input and evaluation of results

Figure 1 provides an example of user input questions for the SWAP Ebola algorithm and the test outbreak details are given below. As shown below the information provided in the table can be applied by a SWAP user to reproduce the SWAP output LANL analysts obtained.

	Guinea 2015	Liberia 2015	Sierra Leone 2015	Guinea 2015
Location	Guinea	Monrovia, Liberia	Sierra Leone	Guinea
Population at risk	12 million	400,000	6 million	12 million
Physician density	0.1	0.014	.022	0.1
History of Ebola	yes	yes	yes	yes
Contact with bush meat or bats	no	No	no	no
case number and time information	5 one day, 27 cumulative (time period not given Tried 5 in one day and 27 in 2 weeks and 1 month	6 case in two months, three case in past week	12 in 1 week and 45 in 1 month (answers on the website	13 in 1 week 47 in one month (extended data is available on the link)
link	http://www.reuters.com/article/2015/05/15/us-health-ebola-guinea-idUSKBN0OO1M22015051 5	http://www.usatoday.com/story/news/world/2015/07/15/ebola-liberia-epidemic-sierra-leone-guinea/30183429/ http://apps.who.int/ebola/current-situation/ebola-situation-report-5-august-2015	http://apps.who.int/ebola/current-situation/ebola-situation-report-5-august-2015	http://apps.who.int/ebola/current-situation/ebola-situation-report-5-august-2015
Range for top 5 matches	69 -62% for 1 day and 2 weeks, 79-67% for 1 months	79-66%	73-58%	1 week 72 -59% 1 month – 72 -67%
Number of expert acceptable matches in top 5	Top 1	Possibly 4 (2nd one is not correct)	Top 1	13 in 1 week – Top 1 47 in 1 month – Only one from the top 5. Both scenarios match to Sudan 1979.
notes	This is a country that is recovering from epidemic.	Post outbreak recovery period	Others cumulative case count too high. possible cut off at 70%	Cumulative case count too high for non-matching outbreaks

Figure 1: A snap shot of user input requirements and the test outbreak table for Ebola

SWAP output can then be compared to data from a test scenario if the information is available (Figure 2). In the comparison given below the 2015 Ebola outbreak in Sierra Leone is expected to have 30-60 cases and duration 8-10 weeks. The top match (74%) is the 1979 Sudan outbreak with 35 cases and duration of two months. The next two outbreaks have very high case numbers compared to 2015 scenario. This type of data allowed at us to determine a lower limit for a relevant similarity score for each disease.
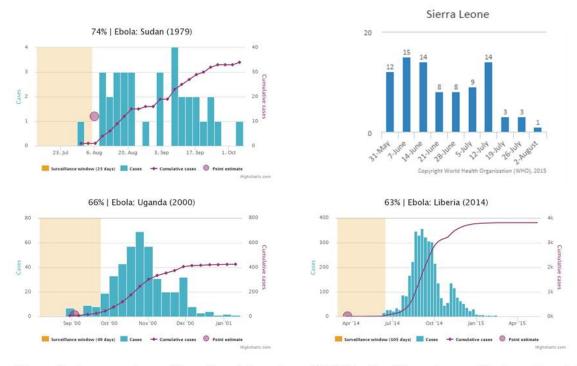
Figure 2: A comparison of top 3 matches from SWAP to the Sierra Leone Ebola outbreak in 2015

Table 1 contains a comparison of properties of the Boston Chipotle outbreak caused by norovirus in late 2015 and top matching SWAP historical outbreaks. The SWAP evaluations were conducted during day 3, day 5 and day 7 of this unfolding outbreak and range of similarity scores for this scenario ranged from 92-75%. The data shows that case count and duration for this outbreak and the top matching historical outbreaks are very similar. These results indicate that SWAP evaluations can be performed as soon as some information on case count, time and properties are available. Information obtained can then be used to enhance and accelerate the outbreak investigation.

	Case count	Time frame	Source of contamination	Product or site event	season
Boston Chipotle	**~150**	**~10 days**	**Not known**	**Site/event**	**winter**
China 2010	427	10 days	Water (norovirus detected from terminal point of water delivery system	Site (highly localized)	winter
Israel 1999	164	5 days	Salad vegetables served at a mess	Site (occurred in	winter

4

			hall + secondary person-to-person	a military base)	
USA Texas 1998	125	10days	Cooked food prepared by an ill person and + soda fountain area	Site (occurred in military base)	autumn
Netherlands 2009	130	12 days	water	Site (students at a recreational fountain)	summer

Table1: Comparison of properties for Boston chipotle outbreaks and top matching SWAP outbreaks. All the outbreaks were caused by norovirus.
The following tables contain data used from test outbreaks for each disease in the SWAP. A SWAP user can use these tables to run the SWAP.

Disease – Chikungunya

	Montpellier, France 2014	Guangdong, China 2010	Brazil, 2015	Florida, USA as of May 5, 2015	French Polynesia, 2015
Location	Montpellier, France	Guangdong, China	Brazil	Florida, USA	French Polynesia
Disease status	Non-endemic	Non-endemic	Non-endemic	Non-endemic	Non-endemic
Population at Risk	30,000-300,000	79 million	200 million	20 million	300,000
Population susceptibility	Naive	Naive	Naive	Naive	Not naive
Human development index	>0.8	0.719	0.744	0.914	Less than 0.8
Outbreak timing	Not summer	Not summer	Summer	Not summer	Summer
Case number	8	20	114 confirmed 6000 suspected	21	70,000 cases
Duration information	2015-07-01 to 2015-07-26	2015-07-01 to 2015-07-15	20-25 week time line	5 months	6 months
link	http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=21108	http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3309566/	Please search the title. It downloads a pdf file. **Number of Reported Cases of Country or Territory 2015 (to week noted) Epidemiological Week / EW**	http://www.cdc.gov/chikungunya/pdfs/2015Table1-050515.pdf	http://www.radionz.co.nz/international/pacific-news/269097/french-polynesia-chikungunya-outbreak-over
Range for top 5	87-69%	98-64%	75-61%	92-63%	67-51%

matches					
Number of expert acceptable matches in top 5	Top 4 Top one is a very good match.	Top 3	Top 4	Top 2	The 1st one and 4th one. Reason for top 1& 4being better match - location characteristic have higher match to test scenario
Notes		Cut off at 70%		Cut off at 70%	

Recommended lower limit for a relevant match – 70%

Disease – Cholera

	Nigeria 2015	Myanmar 2014	Malawi 09/10	PNG09-10	Haiti 2015
location	Bayelsa, Nigeria	Myanmar, south Okkalapa township	Malawi	Papua New Guinea	Haiti
Disease Status	Endemic	endemic	endemic	Non-endemic	endemic
Population at risk	1 million	200,000	16 million	7 million	10million
Case number	23	234/380 tested positive	250	250	4000
Duration information	December 24 – January 28	September 27 – October 8	1 months	3 months	1month (beginning of the outbreak)
link	http://allafrica.com/stories/201501281403.html	http://www.mmtimes.com/index.php/national-news/11951-hundreds-infected-in-cholera-outbreak.html	http://www.who.int/cholera/countries/MalawiCountryProfile2010.pdf?ua=1	http://www.who.int/cholera/countries/PapuaNewGuineaCountryProfile2010.pdf?ua=1	http://www.haitilibre.com/en/news-13483-haiti-cholera-20-communes-on-red-alert.html
Range for top 5 matches	94 -78	90-64	83-77%	71-68%	72-67%
Number of expert acceptable matches in top 5	Top 5	Top 2	Top 4 Last one case count a bit too high	Top 2	Top 3
Notes	Only beginning of outbreak info is	Cut off value 75%		Cut off 70%	Cut off 70%

	available				

Recommended cut off match – 70%

Disease – Dengue

	Singapore 2015	Tamil Nadu, India 2015	Cambodia 2015	Belize 2015	Brazil 2014
Location	Singapore	TN, India	Cambodia	Belize	Brazil
Max pptn	11 in	1	5.9	4	4
Case number	1463, 99 in 7th week	95	2688	110	200,000
Duration information	7 weeks	1 month	6 months	16 weeks	15 weeks
link	https://www.moh.gov.sg/content/dam/moh_web/Statistics/Infectious_Diseases_Bulletin/2015/February/2015_week_07.pdf http://www.todayonline.com/singapore/53-year-old-woman-first-dengue-death-spore-year	http://zeenews.india.com/news/health/health-news/9-die-of-dengue-in-tirunelveli-zone-this-month_1537402.html	http://www.khmertimeskh.com/news/13682/cambodia-reports-sharp-rise-in-dengue-fever-cases-in-1st-half-of-2015/	http://www.paho.org/hq/index.php?option=com_topics&view=article&id=1&Itemid=40734 (click on 2015)	http://portalsaude.saude.gov.br/images/pdf/2014/maio/29/BE-2014-45--4--semana-15.pdf
Range for top 5 matches	98-78%	66-44%	84-61%	71-40%	66-48%
Number of expert acceptable matches in top 5	Top 4 Last one cumulative case count is too high.	None	Top 3 Last 2 cumulative case count too high.	Top 1	None. This is a second year high rate scenario
Notes	Possible cut off at 80%	There is flooding in TN currently, an unusual	Possible cut off at 70%	Possible cut off at 70%	

		situation. Changing rainfall changes matches a bit			

Recommended cut off match – 70%

Disease Ebola

	Guinea 2015	Liberia 2015	Sierra Leone 2015	Guinea 2015
Location	Guinea	Monrovia, Liberia	Sierra Leone	Guinea
Population at risk	12 million	400,000	6 million	12 million
Physician density	0.1	0.014	.022	0.1
History of Ebola	yes	yes	yes	yes
Contact with bush meat or bats	no	No	no	no
Case number and duration information	5 one day, 27 cumulative (time period not given Tried 5 in one day and 27 in 2 weeks and 1 month	6 case in two months, three case in past week	12 in 1 week and 45 in 1 month (answers on the website	13 in 1 week 47 in one mont (extended data available on th link)
link	http://www.reuters.com/article/2015/05/15/us-health-ebola-guinea-idUSKBN0O01M220150515	http://www.usatoday.com/story/news/world/2015/07/15/ebola-liberia-epidemic-sierra-leone-guinea/30183429/ http://apps.who.int/ebola/current-situation/ebola-situation-report-5-august-2015	http://apps.who.int/ebola/current-situation/ebola-situation-report-5-august-2015	http://apps.wh t/ebola/curren situation/ebola situation-repor august-2015
Range for top 5 matches	69 -62% for 1 day and 2 weeks , 79-67% for 1 months	79-66%	73-58%	1 week 72 -59% 1 month – 72 -6
Number of expert acceptable matches in top 5	Top 1	Possibly 4 (2nd one is not correct)	Top 1 Others cumulative case count too high.	13 in 1 week – Top1 47 in 1 month - Only one from top 5. Both scenarios matc Sudan 1979.

				Cumulative cas count too high non-matching outbreaks
Notes	This is a country that is recovering from epidemic.	Post outbreak recovery period	possible cut off at 70%	

I am not sure about a match cut off.

Disease FMD

	Jinning township, Kinmen county, China 2015	Bahrain 2015	Israel 2013	Botswana 2015
Location	Jinning township, Kinmen county, China	Bahrain	Israel	Botswana
Animal identifier	cattle	cattle	cattle	cattle
Cattle density	High (176)	Low (I think)	low	high
Swine density	low	low	low	low
Infected premises	1 case 176 susceptible	13	1 (1 village)	1
Duration information	1 month April 13 –May 1	Feb23 to March 16	1 week	26/07/2015-31/07/2015
links	http://www.oie.int/wahis_2/public/wahid.php/Reviewreport/Review?pop=1&reportid=17682	http://www.gulf-daily-news.com/NewsDetails.aspx?storyid=398212	http://www.oie.int/wahis_2/public/wahid.php/Reviewreport/Review?page_refer=MapFullEventReport&reportid=14420	http://www.oie.int/wahis_2/public/wahid.php/Reviewreport/Review?page_refer=MapFullEventReport&reportid=18268
Range for top 5 matches	77-65%	92-67%	96-75%	96-71%
Number of reasonable matches in top	Top 5	Top 5	Top 5	Top 5

5				
Notes				

Acceptable to have a 70% match cut off.

Disease family _ Gastroenteritis

pathogen	E. coli	E.coli	Salmonella	Salmonella	norovirus	Salmonella
Location	USA 2014	USA 2013	UK 2014	UK		
Source of contamination	Raw clover	Cooked food	eggs	Watermelon from Brazil	Frozen strawberries	Person to person
Product or site/event	product	product	product	product	product	Site/event
season	Spring	winter	summer	December	Autumn	Autumn
military	No	No	No		No	
Case number	19 for all (10 for till May 10th	11, 35 total	55	22	40 (300 in 10 days)	14
Duration information	April 29 – May 20	Dec 30 –Jan 30 (last case April 10	May 21-June 30	Oct30-Nov 30th	3 days	1 week
link	http://www.cdc.gov/ecoli/2014/O121-05-14/index.html	http://www.cdc.gov/ecoli/2013/O121-03-13/index.html	http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=21098	http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20866	http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20719	http://www.clickondetroit.com/news/14-patients-diagnosed-with-salmonella-at-henry-ford-hospital/35566506
Range for top 5 matches	96-78%	97-78%	92-72%	86-77%	77-73% for 3 day 79-68% for 10 day	98-83%
Number of expert acceptable	Top 4. Last one case count too	4matches. One and 4 on the mark. Last one	All 5	All 5	All 5 if using 10	Possibly 4, though not the

	matches in top 5	high.	too high case count			days value	top 1 match
Notes		Identified E.coli as pathogen. Possible cut off at 80%	Identified pathogen as E.coli. Possible cut off at 80%	Pathogen identified as Salmonella	Pathogen – Salmonella or E.coli	Pathogen narrowed to E.coli or norovirus but needed 10 day data	Pathogen identified wrongly as Norovirus or Shigella. We don't have hospital as source at the moment

Applying this to real life situation - Boston Chipotle outbreak

70-80% match required for pathogen identification

80 over the weekend (Dec 5-7)

	outbreak	% match
1	Norovirus: Israel 1999	92%
2	Campylobacteriosis: Belgium 2010	80%
3	Norovirus: New Jersey 2006	79%
4	Norovirus: China 2010	79%
5	Norovirus: Netherlands 2009	76%

Stool tests for Norovirus and *E.coli* were in progress on December 8th morning. Early lab tests reported presence of norovirus on Wednesday December 9th. On December 11th lab results confirmed the absence of other bacterial pathogen (Boston public health commission).

100 cases by Wednesday (Dec 9th)

	outbreak	% match
1	Norovirus: China 2010	91%
2	Norovirus: Israel 1999	79%
3	Campylobacteriosis: Finland 2012	75%
4	Campylobacteriosis: Spain 2010	75%
5	Norovirus: Texas 1998	73%

Note:
Norovirus China 2010, Netherlands-2009 and campylobacteriosis outbreaks were caused by water based contamination.

141 cases by Friday (Dec 11)

	outbreak	% match
1	Norovirus: China 2010	88%
2	Norovirus: Israel 1999	77%
3	Campylobacteriosis: Finland 2012	76%
4	Norovirus: Netherlands 2004	76%
5	Campylobacteriosis: Belgium 2010	75%

Disease - influenza

ID	Argentina 2009	New Zealand 2009	Cambodia 2013	Zhejiang, China 2013	United States 2011
Location	Argentina	New Zealand	Cambodia	Zhejiang, China	United States
Strain	H1N1	H1N1	H5N1	H7N9	H3N2v
Person to person	Yes	Yes	no	Yes	no
Case number	7,173 confirmed (811,940 Influenza-like cases)	3086 confirmed	26	40	12
Duration information	May 16 – August 23 2009	April 28 – August 21 2009	February 1 – Dec 10 2013	March 9 – April 17 2013	August 17 – December 23 2011
link	http://www.flu.gov/pandemic/global/final.pdf	http://www.flu.gov/pandemic/global/final.pdf	http://www.who.int/influenza/human_animal_interface/H5N1_cumulative_table_archives/en/	http://www.who.int/influenza/human_animal_interface/influenza_h7n9/ChinaH7N9JointMissionReport2013u.pdf?ua=1	http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6051a4.htm
Range for top 5 matches	75-68%	83-71%	100-81%	80-71%	80-77%
Number of expert acceptable matches in top 5	Top 4	Top 5	Top 5	Top 5	Top 5
Notes	They are all H1N1		All matches to novel influenza	All matches to novel influenza	All matches to novel influenza

Note: we only have 5 seasonal influenza

Cut off match of 70% is reasonable.

Disease – Leptospirosis

	Maharashtra, India 2015	Florida, USA 2005	Manila, Philippines 2009	Karnataka, India 2011
Location	Maharashtra	Florida	Manila	Karnataka
HDI	medium	Very high	medium	medium
Contact with water	Heavy rain fall	swimming	rain	yes
Flooding or population displacement	yes	no	yes	no
Case number	130	44 in 1 month 12 in 3 weeks	300	100 (estimated peak) total cases in 2 weeks 286
Duration information	Start of monsoon (June) to mid-September		1 October to Nov 19	4 days
link	http://www.dnaindia.com/mumbai/report-19-die-due-to-leptospirosis-this-year-2125261	http://cid.oxfordjournals.org/content/50/6/843.full.pdf+html	http://reliefweb.int/sites/reliefweb.int/files/resources/D87F541373D245BC4925767A000C5761-Full_Report.pdf http://www.ncbi.nlm.nih.gov/pmc/articles/PMC33	http://www.scopus.com/record/display.url?eid=2-s2.0-84930817082&origin=inward&txGid=8F21F1C0A98CEB873852A4C8F44CD289.aqHV

			10081/	0EoE4xlIF3hgVWgA%3a1
Range for top 5 matches	67-57%	73-66%	88-66%	73-69%
Number of expert acceptable matches in top 5	Top 3.	Top 3	Top 4 One of the 66% is OK, but other is not due to high case number.	Top 3.
Notes	Cut off 65%	cut off at 70%		Cut off at 70%

Cut off match of 70% is reasonable.

Disease – Malaria

	Puerto Rico USA 2015	**Venezuela 2014**	**Madagascar 2015**	**Goa, India 2015**
location	Puerto Rico, USA	Venezuela 201	Madagascar	Goa
Population at risk	4 million	31 million	24 million	1.8 million
Cases per thousand	Less than 1	More than 1	More than 1	Less than 1
Case number	3	1000 (week 44 has answers)	200, 000	650
Duration information	One month	Week 1 of 2014	5 months	4 months
link	http://promedmail.org/direct.php?id=20150720.3524406	http://www.mpps.gob.ve/index.php?option=com_phocadownload&view=category&id=43:ano2014&Itemid=915	http://www.lexpressmada.com/blog/actualites/paludisme-le-bilan-salourdit-36369/	http://timesofindia.indiatimes.com/city/goa/Mapusa-witnessing-an-increase-in-malaria-cases/articleshow/47650415.cms
Range for top 5 matches	80-76%	61-53%	75-42%	75-66%
Number of expert acceptable matches in top 5	1st two 80&78% matches have case count in 10s. Next three has case count in 100s. All 5 matches are reasonable.	All 5 have the potential to be correct. 1st two are 60K last three 600K range	1st two cut off at 50%	All 5 seems reasonable
Notes	Both sets of scenarios are reasonable outcomes based on control measures	This is very bad outbreak with probably cumulative		

	implemented	case 100-200K		

The matches for these test outbreaks are over all a little low. The library is too inclusive, so getting test outbreaks was difficult.

Disease – Measles

ID	Denmark 2013	Denmark 2014	Bosnia and Herzegovina 2014-2015	Congo 2015
Population at risk	5.6 million	5.6 million	3.8 million	71 million, Kataga 5 million
Country immunization %	90	90	89	77
Affected region %	90	90	89	77 (kataga 51%)
Case number	9	18	106	20,000
Duration information	5 weeks	8 weeks	5 weeks	6 months
link	http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=21254	http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=21254	http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=21047	http://www.msf.org/article/drc-katanga-measles-epidemic-keeps-worsening
Range for top 5 matches	95-85%	88-79%	94-81%	85-66%
Number of expert acceptable matches in top 5	Top 1. Case numbers a bit too high for 3 out of 5.	Top 1.4 out of 5 outbreaks same but slightly different order for Denmark 2013 and 2014 tests	Top 5 #1 is a very good match for case count as well as time	Top 4
Notes				Cut off at 70%

Cut off match of 70% is reasonable.

Disease - Meningococcal

ID	UK 2000	Sudan 2006	Nigeria 2015	USA 2013-14	Niger 2015
Location	United Kingdom	Sudan	Nigeria	Princeton University, US	Niger
Disease status	Non-endemic	Non-endemic	Endemic	Non-endemic	Endemic
strain	W and A	A	C	B	C
Fatality	18.2%	6.9%	7.7%	11.1%	6.9%
Case number	22	231	652	9	5855
Duration information	March 21 – April 11 2000	September 1 -November 8 2006	January 26 – March 5 2015	March 2013 – March 2014	January 1 – May 10 2015
link	http://www.who.int/csr/don/2000_04_11/en/ ; http://www.who.int/csr/don/2000_04_21b/en/	http://www.who.int/csr/don/2006_11_21/en/	http://www.who.int/csr/don/13-march-2015-nigeria/en/	http://www.nfid.org/idinfo/meningitis/meningococcal-b-college-outbreaks.html	http://www.who.int/mediacentre/news/situation-assessments/meningitis-niger/en/
Range for top 5 matches	79-41%	71-64%	71-64%	69-31%	74-60%
Number of expert acceptable matches in top 5	Top 2	Top 5	Top 2	Top 2	4 and 5 seem to be best match for case count and time
Notes	Cut off at 70%			Cut off at 60%	Lower score due to strain not matching.

Cut off match of 60% is reasonable.

Disease –Plague

ID	USA 2015	Madagascar 2015	Zambia 2015
location	Santa Fe	Madagascar	Zambia
Mortality	25%, low	high	low
Plague type	septicemic	pneumonic	bubonic
season	fall	Rainy season	April in Zambia, Rainy season
Case number	2 in Santa Fe, 4 in NM	14 cases, 10 death	13 cases, 3 deaths
Duration informatio n	3 months or 1 month	17 Aug-27 Aug	3 months or 1 month
link	http://nmhealth.org/news/disease/2015/9/?view=313	http://who.int/csr/don/06-september-2015-plague/en/	http://outbreaknewstoday.com/plague-outbreak-kills-three-in-zambia-24623/
Range for top 5 matches	Santa Fe -50-44% NM score 63-51%	87-52%	81-53%
Number of expert acceptable matches in top 5	None	Top 3 & 5th The fourth match is very different in terms of case count and time from other 4.	Three months looks better matches
Notes			Possible cut off at 65-70%

Cut off match of 65-70% is reasonable.

Disease Q-fever

ID	**USA 2011**	**UK , Scotland 2006**	**Baranya, Hungary 2013**	**Bavaria, Germany 2014**
Location	USA (WA, MT)	UK , Scotland	Baranya, Hungary	Bavaria, Germany
Pop. at risk	~100,000	300 (meat plant)	396,633	12.6 million
Farm/herd proximity	<5km	NO (questionable)	<5km	5km-20km
season	summer	summer	Summer	Spring
Case number	11	24	70	12
Duration information	June - July	1 month	April –July 2013	Jan – April 2014
link	http://www.capitalpress.com/content/mw-Q-fever-070111	http://www.documents.hps.scot.nhs.uk/ewr/pdf2006/0629.pdf	http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20863	http://promedmail.org/direct.php?id=20140511.2464862
Range for top 5 matches	75-62%	92-64%	82-63%	88-51%
Number of expert acceptable matches in top 5	Top 2. Case count seems too high for all, but time is good match for top 2.	Top 5	Top 1 matches both case count and time frame. Others – either case count or time doesn't match. Historic epidemic predicts shorter duration.	Top 3

Notes				70% cut off most likely fits best. More cases are expected

Cut off match of 70% is reasonable.

Disease - Tularemia

	USA 2015	USA 2015	Spain 2014	Russia 2014
Location	South Dakota, US	Colorado, US	Spain	Russia
Season	summer	summer	summer	Late summer
Endemicity and rural vs. urban	Endemic, rural	Endemic, rural	Endemic, rural	Endemic, rural
Type			Ulceroglandular and respiratory	
Case number	7	16	12	7
Duration informatio n	June-July 2015	June 2015	August 2014 (1 month)1	Sept. 2014
link	http://outbreak newstoday.com /tularemia-sickens-7-in-south-dakota-since-june-30075/	http://outbreakne wstoday.com/color ado-tularemia-5-cases-reported-in-weld-county-89951/	http://promedmail .org/direct.php?id= 20140815.268468 9	http://outbreaknews today.com/tularemia -outbreak-reported-in-russia-34925/
Range for top 5 matches	82-63%	75-66%	90-69%	76-63%
Number of expert acceptable matches in top 5	Top 2	Top 2	Top 1 The first one looks very good. Others seem to fall to two categories with low or higher case count. Both might be possible	Not sure. Case counts again fall into two categories
Notes	70% cut off	Cut off 70%		

I tried to characterize tularemia type based on matches, but the library is too limited to achieve this. 70% cut off seem reasonable in some test scenarios. Others it is too difficult to determine.

Disease – West Nile virus

	Greece 2012	**USA 2014**	**New Mexico 2015**	**Ontario 2014**
Location	Greece	Los Angeles, CA, USA	New Mexico	Ontario
Outbreak timing	Not tail end	Tail end	Tail end	Tail end
Rural or urban	urban	urban	rural	rural
Climate, precipitation, temperature	csa	csb	Bsh	Dfc
Case number	18 cs in three weeks	55	11	20
Duration information	80 cases in 3 months	4 weeks	4 months	4 months
link	http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20758	http://www.publichealth.lacounty.gov/acd/docs/West%20Nile/WNVepi2014.pdf	http://www.cdc.gov/westnile/statsmaps/preliminarymapsdata/histatedate.html	http://healthycanadians.gc.ca/diseases-conditions-maladies-affections/disease-maladie/west-nile-nil-occidental/surveillance-eng.php
Range for top 5 matches	95-77%, 91-77%	80-67%	74-48%	68-52%
Number of expert acceptable matches in top 5	Top 5	Top 5	Top 1	Top 1. For others case count is too high
Notes	Four outbreaks are common for the scenarios		70% cut off	

Cut off match of 65-70% is reasonable.