Monitoring Stance towards vaccination in Twitter messages

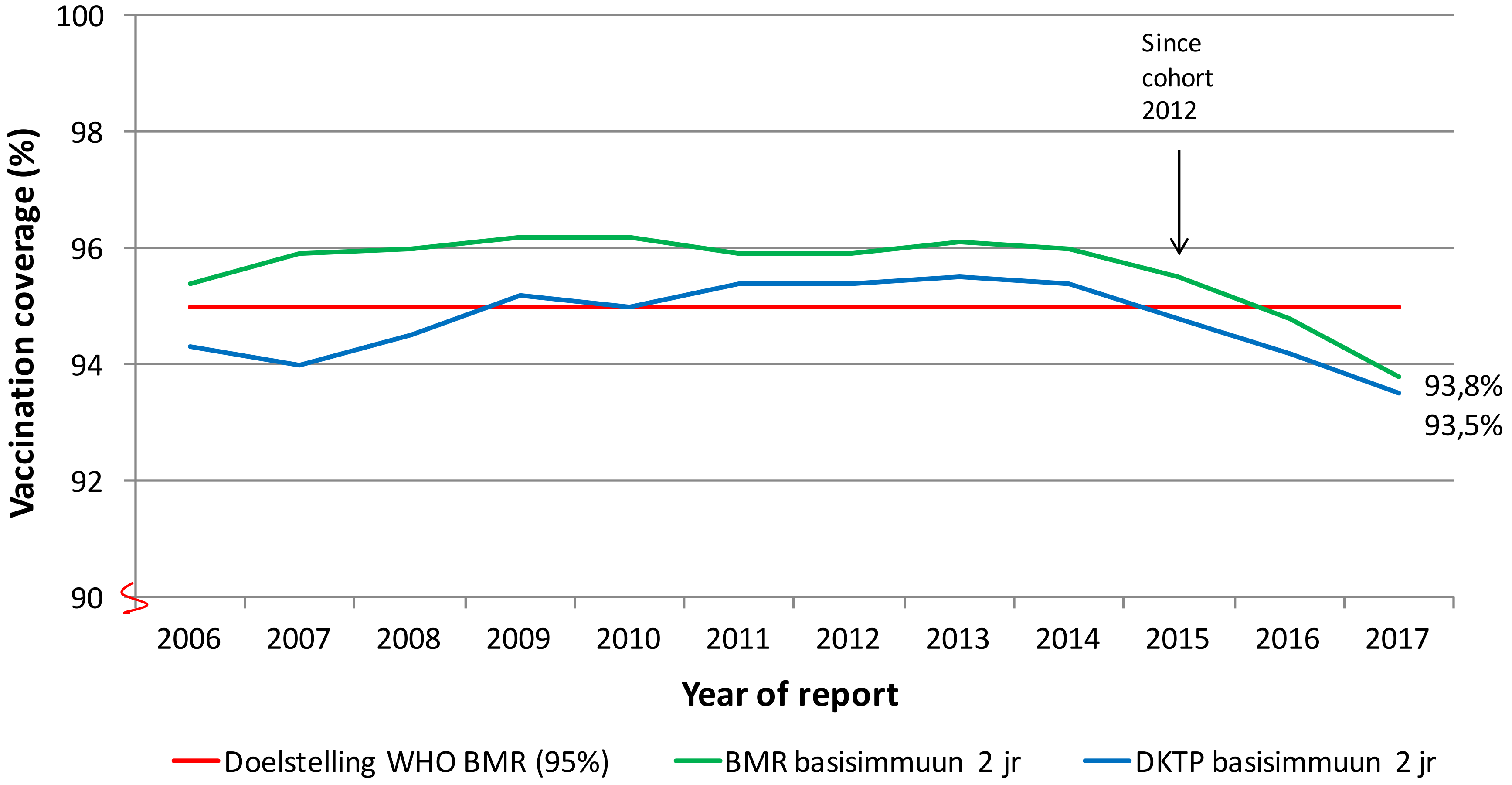
CLIN28, January 26 2018

Florian Kunneman, Liesbeth Mollema, Antal van den Bosch, Mattijs Lambooij, Albert Wong and Hester de Melker

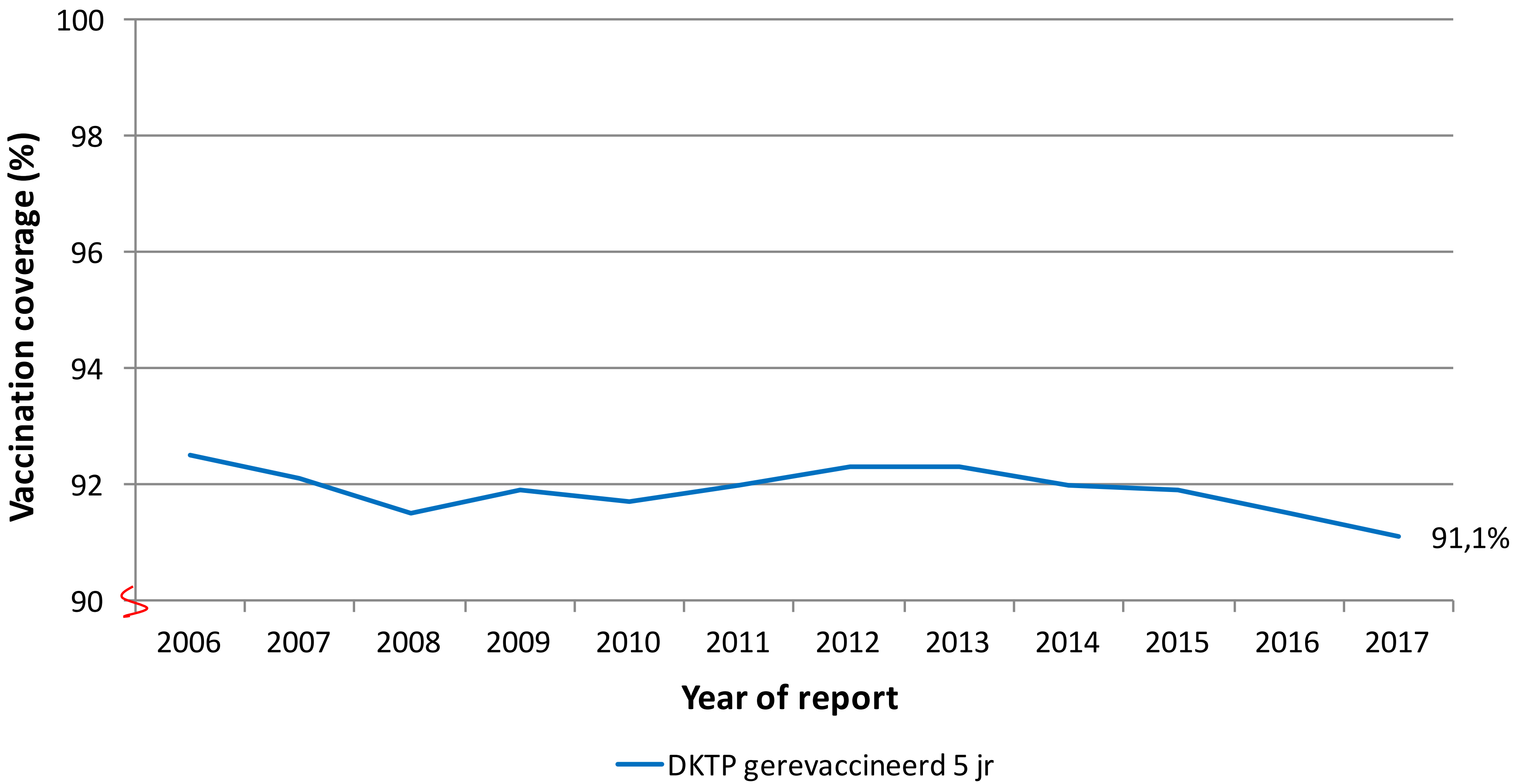
Radboud Universiteit

Rijksinstituut voor Volksgezondheid en Milieu
Ministerie van Volksgezondheid, Welzijn en Sport

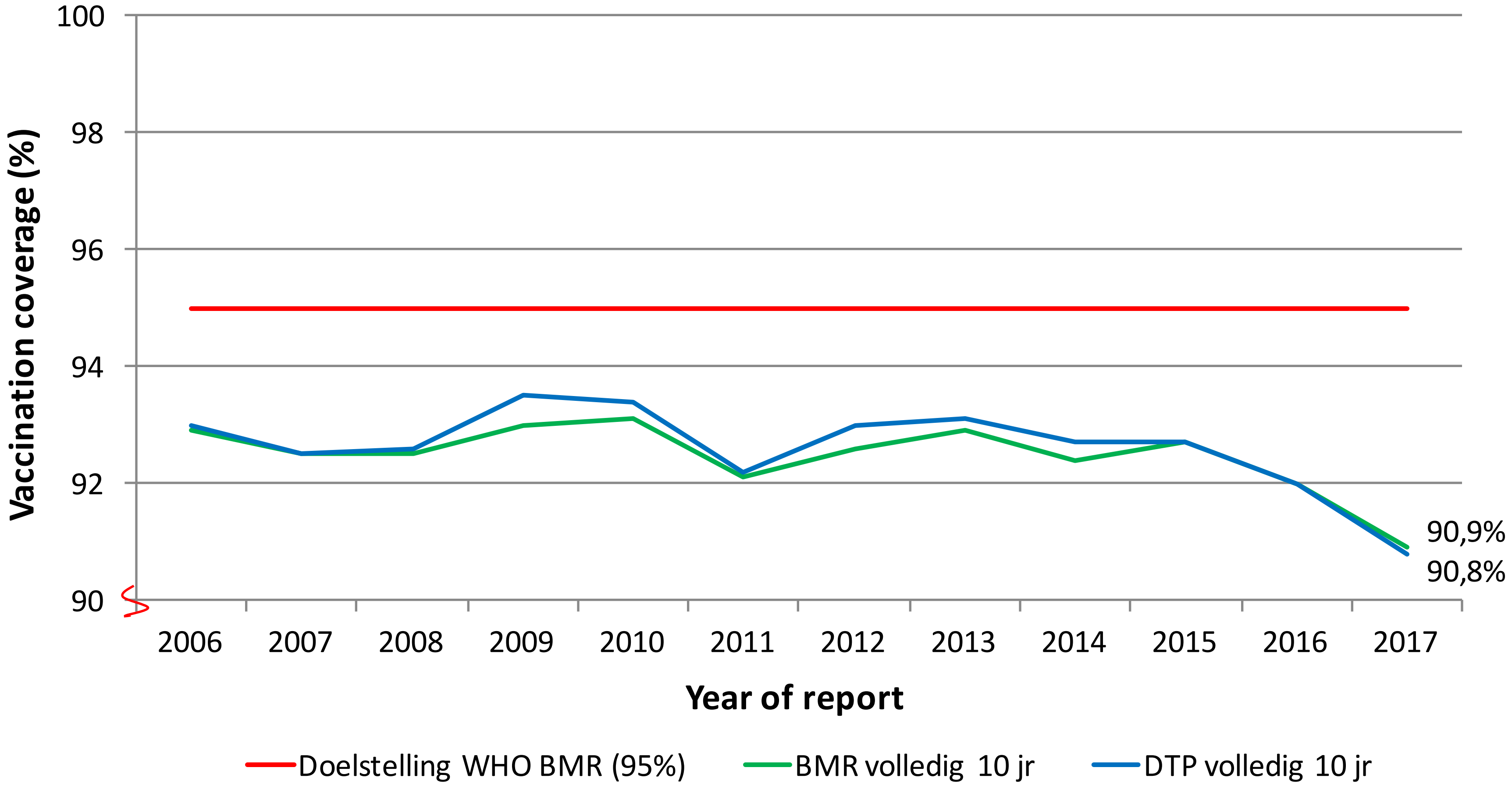
State of affairs vaccination rate baby's



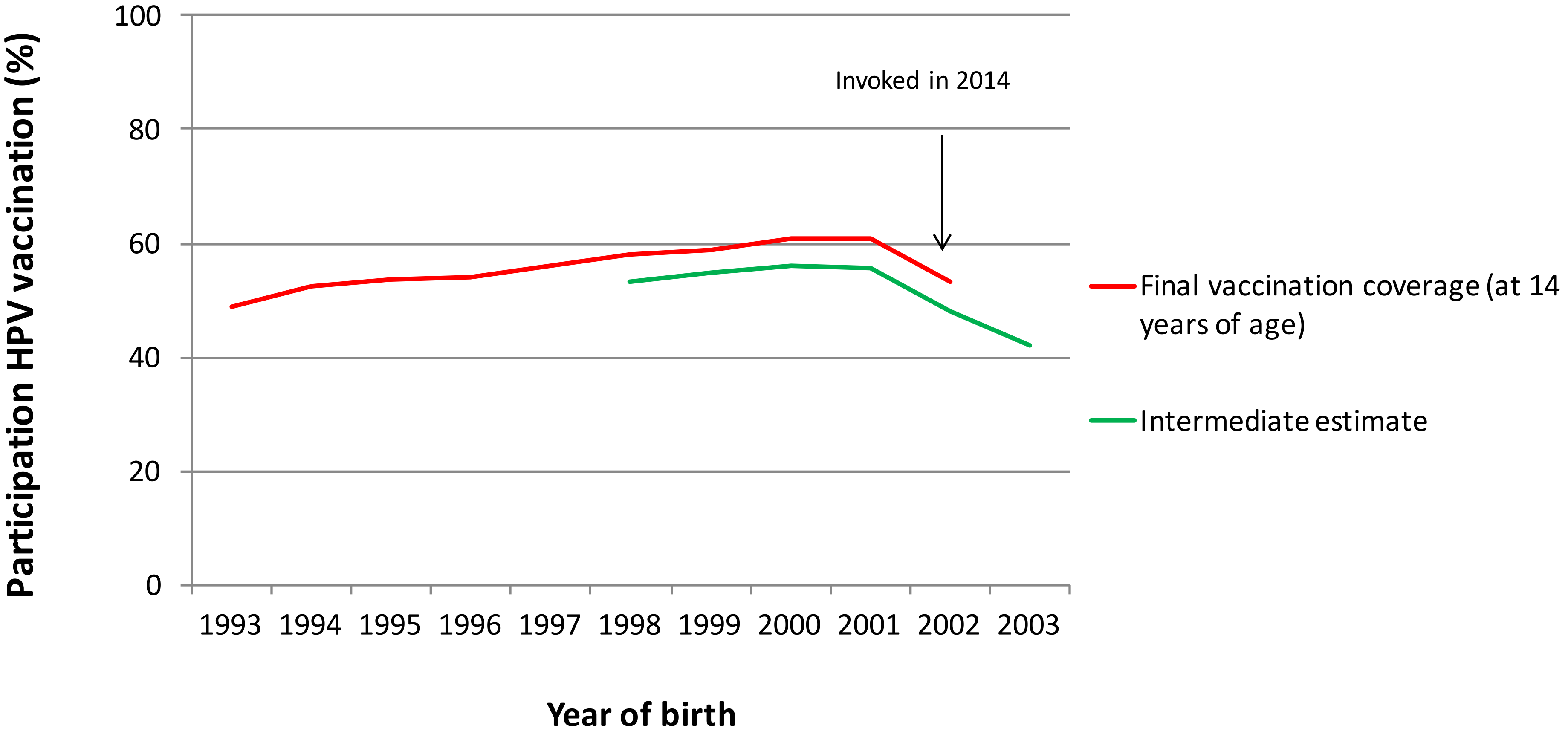
State of affairs vaccination rate toddlers



State of affairs vaccination coverage pupils



Vaccination coverage: first time decrease HPV

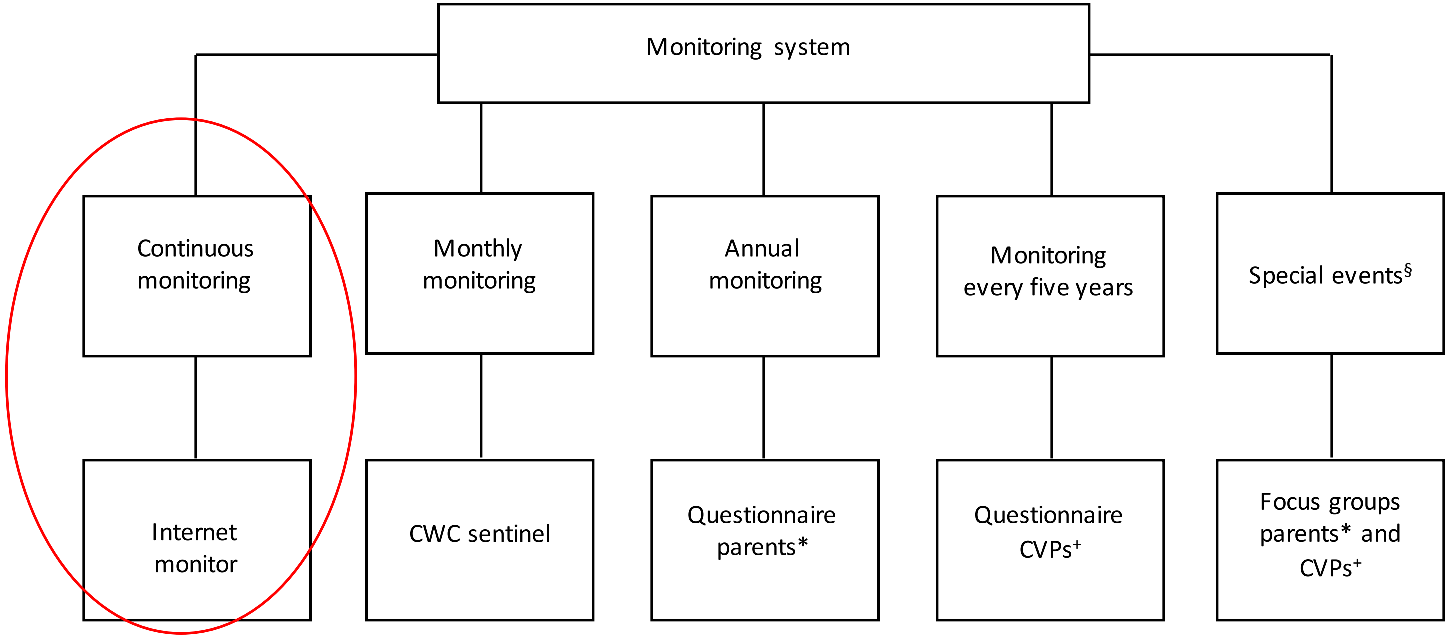


van Lier et al. Vaccination coverage and annual report National Immunisation Programme in the Netherlands 2016. Bilthoven: RIVM; 2017. http://www.rivm.nl/bibliotheek/rapporten/2017-0010.pdf

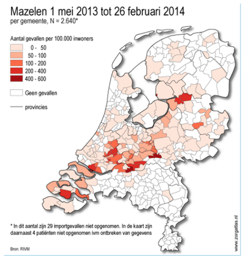
Diversity of reasons for decline

- No fright of diseases
- More attention for side-effects
- Spread of negative stories on the Web
- More critical stance of parents
- Less faith in the government
- ...

Surveillance of vaccination uptake



Online activity “vaccination” from Jan 2012 to Nov 2017



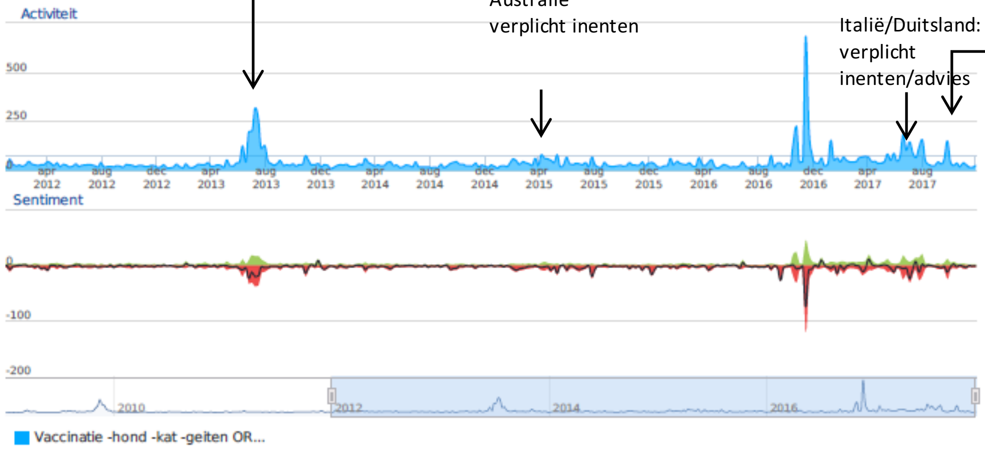
Column Happinez
Lubach
RTL Late Night
Zorgnu.nl
Pauw

Uitbreiding vaccin meningokokken

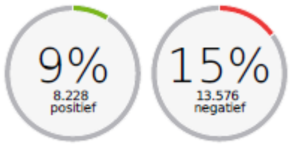
Activiteit & Sentiment

Australië verplicht inenten

Italië/Duitsland: verplicht inenten/advies



93.454
berichten



van January 1, 2012
tot en met November 28, 2017

Trending topics

Bronnen

Type bronnen	Berichten
1. Twitter	51.288
2. Nieuws	13.203
3. Forum	12.181
4. Blog	8.202
5. Facebook	7.033
6. Instagram	470
7. Google+	330
8. Pinterest	233
9. review	226
10. YouTube	145
11. Blendle	56
12. LinkedIn	54

Research aim

- Automatically categorizing vaccination tweets:
  - Relevant - Irrelevant
  - **Stance** (Positive / **Negative** / Neutral / Not clear)
  - Sentiment (Informative / Anger / Concern / Relief / Other)

- Can **negative stance** towards vaccination better be predicted using supervised machine learning than lexicon-based sentiment analysis?

@USER Not unless 2 of your 5 children almost died after a vaccination. Oh well... I gladly sacrifice my children for the herd.

@USER Behalve dan als 2 v je 5 kinderen bijna dood gingen aan n vaccinatie. Maar goed... Ik offer m'n kinderen op voor de grote groep.

RIVM folder: If children get ill right after an inoculation, it seems as though the stab caused it, but that is HARDLY ever the case.

RIVM folder: Als kinderen vlak na een inenting een ziekte krijgen, lijkt het alsof dat door de prik komt, maar dat is BIJNA nooit zo.

People from the bible belt are afraid of getting contaminated with diseases that refugees carry. There own problem, says Schippers, they could have themselves inoculated. HUH

Mensen uit de bijbel belt zijn bang voor de ziektes van vluchtelingen. Tja zegt Schippers moeten ze zich maar laten inenten. HUH

Data

- Focus on Twitter
- Source: TwiNL (twinl.surfsara.nl)
- Query:
  - vaccinatie, vaccin, vaccineren, rijksvaccinatieprogramma, vaccinatieprogramma, intenting, inenten
  - With and without hashtag (‘#’)
  - Januari 1 2012 – February 8 2017

Data

- Number of tweets after query: 96,566
- Pre-filtering:
  - Removal of retweets
  - Removal of tweets with a URL
- Number of tweets after pre-filtering: 23,836

Manual annotations

- Automatically categorizing vaccination tweets:
  - Relevant - Irrelevant
  - **Stance** (Positive / **Negative** / Neutral / Not clear)
  - Sentiment (Informative / Anger / Concern / Relief / Other)

- ‘if the author is negative about the act of vaccination (or downplays the gravity of a disease which is vaccinated against), label the tweet as <u>‘negative’</u>

Manual annotations

- Online annotation tool
- 65 coders
- Preparation:
  - Annotation manual
  - Training round with feedback for 15 tweets

Manual annotations

Statistics

Number of annotations	14918,0
Number of annotated tweets	8259,0
Annotated twice	6540,0
Annotated once	1719,0
Number of annotators	65,0
Avg. number of annotations by annotator	229,5
Standard deviation	444,8
Median	122,0
Most active annotator	2388,0

Manual annotations

- Agreement
  - Relevance
    - Percent agreement : 0.71
    - Cohen's Kappa : 0.23
  - Polarity
    - Percent agreement : 0.54
    - Cohen's Kappa : 0.35
  - Sentiment
    - Percent agreement : 0.54
    - Cohen's Kappa : 0.35

Experiment

Numbers (annotated: 8,259, not annotated: 15,577)

		Strict	Lax	One
Relevance	Relevant	2,249	3,660	1,287
	Irrelevant	636	1,704	435
Polarity	Negative	343	845	191
	Positive	1,312	1,317	523
	Neutral	345	926	278
	Not clear	253	508	286
Polarity and Sentiment	Positive + Information	300	829	213
	Positive + Frustration	392	444	168
	Positive + Concerned / Relieved / Other	620	204	142

Experiment

- Target: Automatic identification of tweets with a negative stance towards vaccination
- Supervised Machine Learning
- Labeled data used for:
  - Trainen ML classifier (strict en/of lax en/of one)
  - Evaluatie ML classifier (strict)
- Features:
  - Word unigrams, bigrams, trigrams
  - Weight: binary
  - Pruning: Most frequent 15,000 features
- Evaluation by 10-fold cross-validation

Experiment

1: Training data

2: Labels

3: Classifiers

SVM (Linear Kernel, C=1, class weight = balanced)

Naive Bayes (alpha=0.0, fit_prior=False)

4: Irrelevance filter

		Strict	Lax	One
Relevance	Relevant	2,249	3,660	1,287
	Irrelevant	636	1,704	435
Polarity	Negative	343	845	191
	Positive	1,312	1,317	523
	Neutral	345	926	278
	Not clear	253	508	286
Polarity and Sentiment	Positive + Information	300	829	213
	Positive + Frustration	392	444	168
	Positive + Concerned / Relieved / Other	620	204	142

Baseline 1: Sentiment Analysis

Pattern

Coosto

Baseline 2: Random

Results

Legend:

Baseline
SVM – Strict + lax
SVM – Strict + lax + one
NB – Strict

	Hierarchical			Flat		
	Precision	Recall	F1	Precision	Recall	F1
Random (0.50)				0.11	0.46	0.18
Random (0.15)				0.12	0.15	0.13
Pattern				0.14	0.34	0.20
Coosto				0.20	0.31	0.25
Negative - Other				0.30	0.38	0.34
Negative - Irrelevant - Other	0.12	0.10	0.11	0.29	0.39	0.34
Negative - Neutral – Positive – Not clear - Irrelevant	0.13	0.12	0.12	0.29	**0.47**	**0.36**
Negative - Neutral – Positive+Frustration – Positive+Information - Positive+Concerned/Relieved/Other - Irrelevant	0.14	0.16	0.15	**0.34**	0.39	**0.36**

Analysis: Pattern vs. ML

		ML	
		Other	Negative
Pattern	Other	1718	372
	Negative	604	192

Both **correctly** Negative : 63 (33 %)
Only ML **correctly** Negative : 99 (44 %)
Only Pattern **correctly** Negative : 51 (8 %)

Analysis: Pattern vs. ML

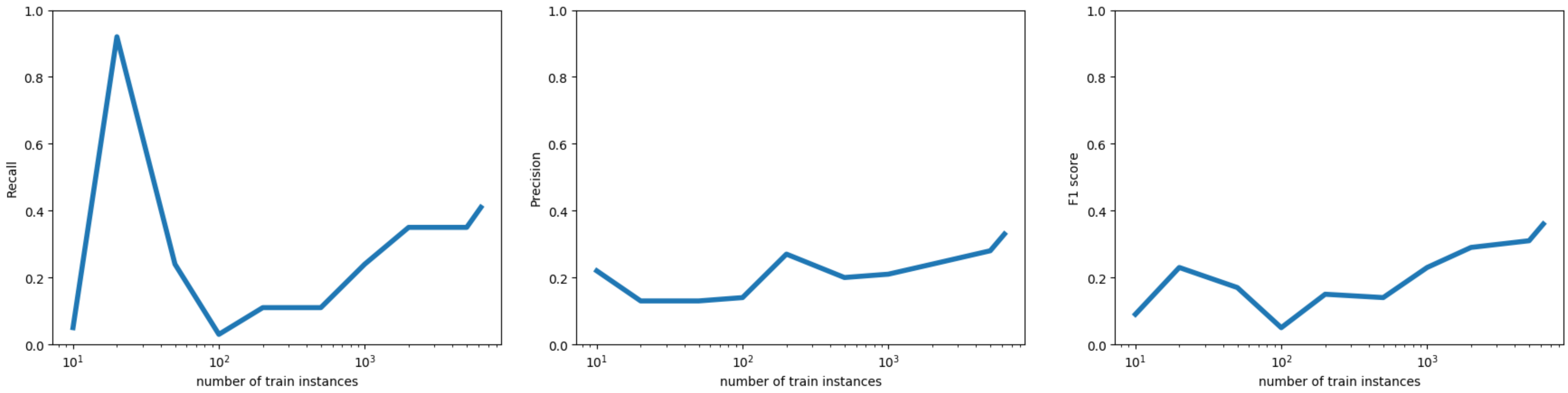
		ML	
		Other	Negative
Pattern	Other	8954	2225
	Negative	3383	1015

Both **correctly** Negative : 36 %
Only ML **correctly** Negative : 33.5 %
Only Pattern **correctly** Negative : 21 %

Analysis: Pattern + ML

Cat	Pr	Re	F1	TPR	FPR	AUC	Tot	Clf	Cor
Other	0.92	0.62	0.74	0.62	0.39	0.62	2543	1718	1585
Negative	0.18	0.61	0.28	0.61	0.38	0.62	343	1168	210
micro	0.83	0.62	0.69	0.62	0.39	0.62	2886	2886	1795

Analysis: Learning curve



Conclusions

- Predicting Negative tweets at F1=0.36 at best
- Comparing:
  - Random baseline: 0.18
  - Sentiment analysis: 0.25
- More training tweets!
  - 8,259Training tweets, 343 reliably labeled negative
  - Active learning? Semi-supervised?
- More reliable training tweets!
  - Maximally two annotations per tweet, low agreement
- More features!
  - Word-to-vec
  - Informed (arguments, …)
- More context!
  - User tweets
  - Conversation history

Questions?

Best ML

Cat	Pr	Re	F1	TPR	FPR	AUC	Tot	Clf	Cor
Irrelevant	0.59	0.27	0.37	0.27	0.05	0.61	633	294	172
Negative	0.29	0.47	0.36	0.47	0.16	0.66	343	564	161
Neutral	0.26	0.34	0.3	0.34	0.13	0.61	345	451	118
positief	0.61	0.63	0.62	0.63	0.33	0.65	1312	1354	832
Not Clear	0.14	0.13	0.13	0.13	0.07	0.53	253	223	32
micro	0.49	0.46	0.45	0.46	0.2	0.63	2886	2886	1315

Pattern

Cat	Pr	Re	F1	TPR	FPR	AUC	Tot	Clf	Cor
negatief	0.14	0.34	0.2	0.34	0.27	0.53	343	796	115
other	0.89	0.73	0.8	0.73	0.66	0.53	2543	2090	1862
micro	0.8	0.69	0.73	0.69	0.62	0.53	2886	2886	1977