Identification of Emerging Key Opinion Leaders Using Bibliometric Data Analysis

Authors: Graber A.[1], Taylor III H.[2], Stapleton A.[2], Boothe K.[1], Mello Jr. A.[2], Rametta M.[1]
[1]Bayer HealthCare; Whippany, NJ, USA
[2]SteepRock Inc.; Washington, CT, USA

BAYER
RUTGERS Institute for Pharmaceutical Industry Fellowships
SteepRock

Background

Since the initial development of the diffusion of innovation model and idea of the Key Opinion Leader (KOL)[1], there has been a great deal of interest in the identification of suitable KOLs for industry involvement. Current market dynamics with increased competition and stricter government regulations drive a need for effective KOL identification and management to support product development and market penetration.[2] Active, sought-after KOLs have limited time to spend on industry initiatives.

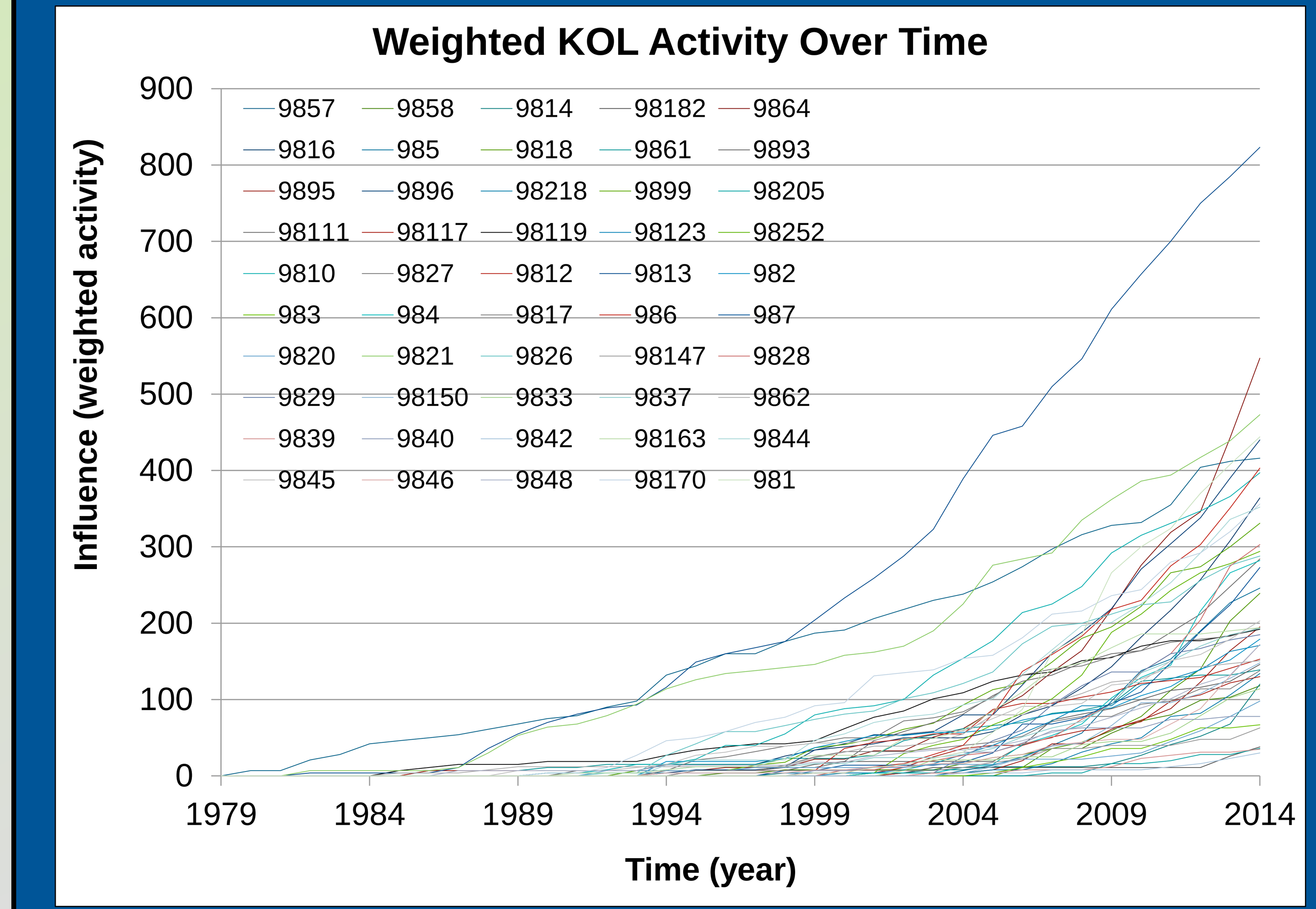
Currently, there are a variety of techniques used to identify KOLs, ranging from subjective and sociometric to objective and bibliometric. Although sociometric techniques play a role in emerging KOL identification, previous studies have shown that questionnaire based approaches to opinion leader identification may lead to unusable information that is highly limited by responder bias.[3] There is limited published information on objective means of emerging KOL identification, especially means built specifically from publicly available data.

Objective

Using publicly available bibliometric activity data as an indicator of influence, this research develops a model for the prediction of Key Opinion Leader (KOL) emergence and demonstrates its use.

Methods - Data Set

- Wet Age-related Macular Degeneration (Wet-AMD)
- 50 KOLs: 1979 – 2014
- Total of 2,877 data points
- Bibliometric Data
  - Publications
  - Trials
  - Grants



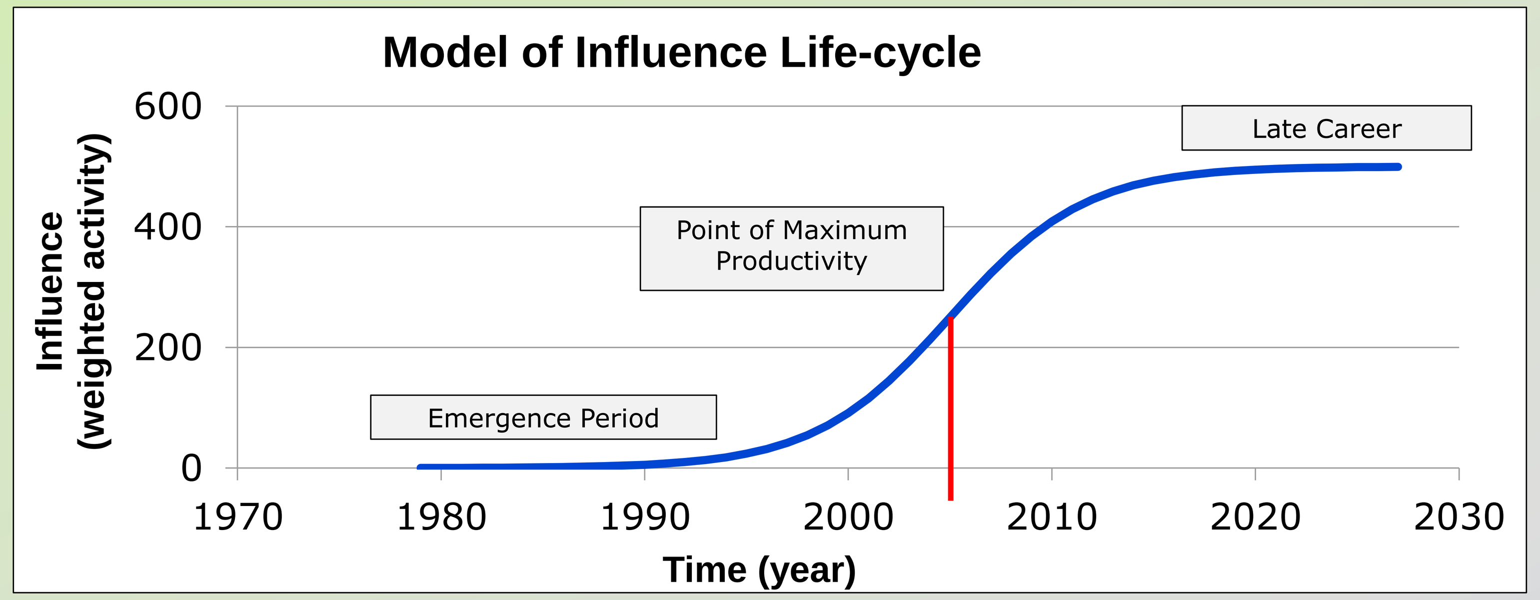
Methods - Model

Fitting the Data:
- Non-linear regression required
- MySQL "R" and Excel were used for model fitting

Logistic Growth Model:
- Functional Form:

$$i(t) = \frac{\theta_1}{1 + e^{-(\theta_2 + \theta_3 t)}}$$

Where:
- $i(t)$ is the influence of the subject over time
- θ_1 is the maximum influence attainable over time by the KOL
- θ_2 is an offset that determines the location of the curve in time
- θ_3 controls the sensitivity of the curve to time



Results – Model Fit

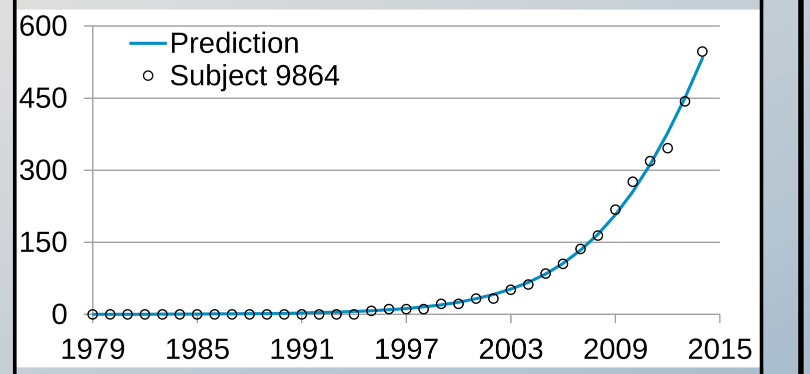
Model Characteristics:
- Well-known and understood (first published by Velhulst in 1845)[4]
- Continuous, Non-Linear Function in *t*
- When parameterized it:
  - Initially grows gradually
  - Reaches peak activity
  - Tapers off as time passes

Matches the profile of influence growth.

Model Features:
- Maximum influence is θ_1
- Time frame of maximum productivity, t_{max}, can be derived analytically:

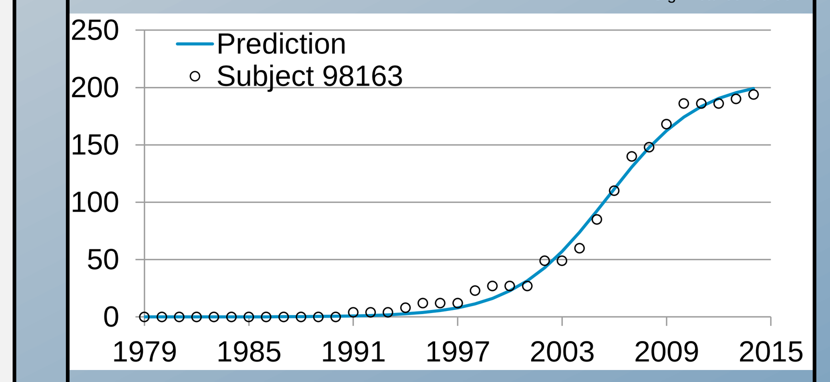
$$t_{max} = \frac{-\theta_2}{\theta_3}$$

Data vs. Model Predictions
2 Samples with Coefficients



	Estimate	Std Err	t Value	Pr (> |t|)
θ_1	1494	292.05	5.1	1.3E-5*
θ_2	-9.26	0.17	-54.1	8.5E-34*
θ_3	0.248	0.01	20.8	1.5E-20*

*Significance < 0.001

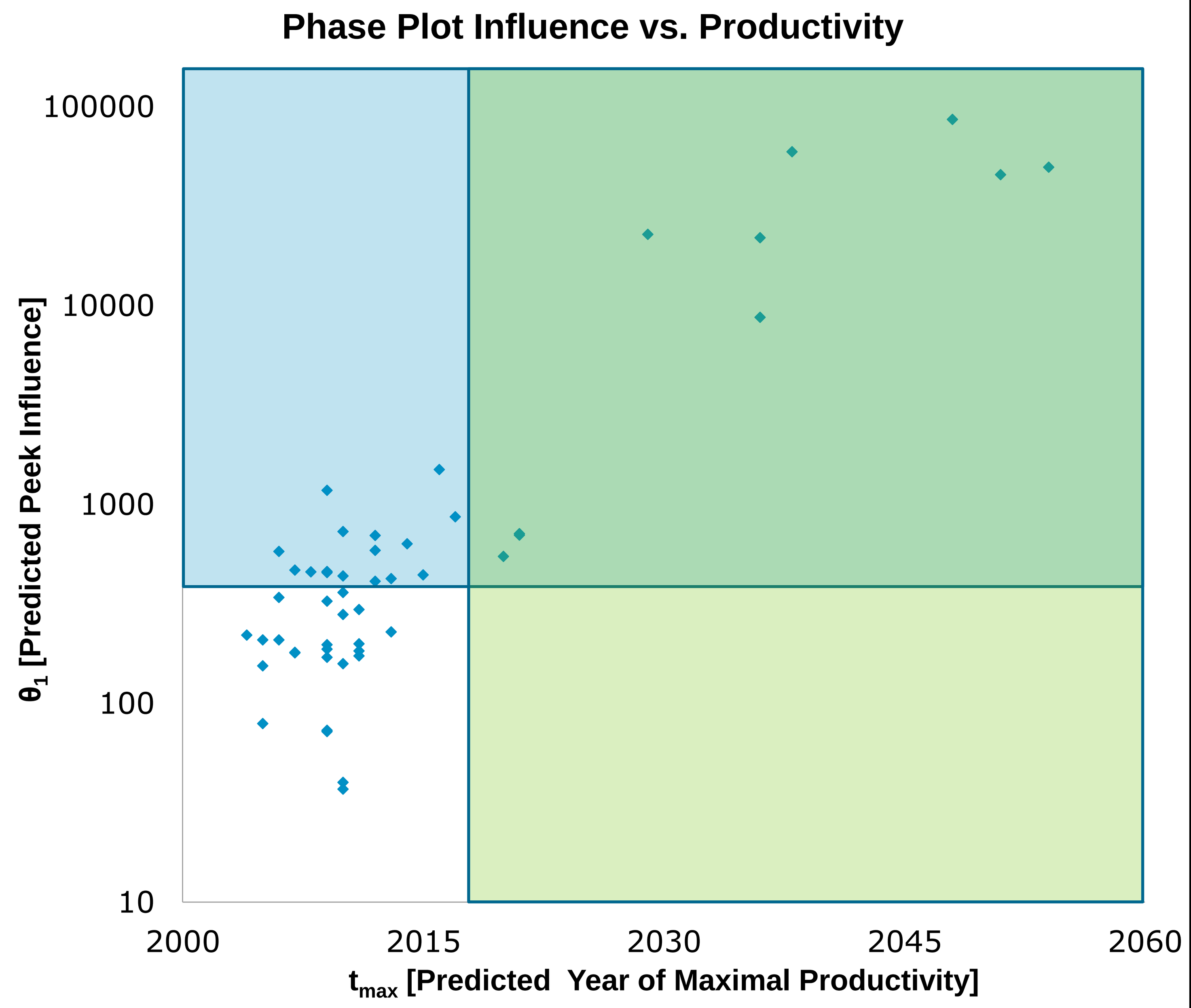


	Estimate	Std Err	t Value	Pr (> |t|)
θ_1	208	5.29	39.3	2.7E-29*
θ_2	-9.95	0.53	-18.7	3.8E-19*
θ_3	0.374	0.02	17.1	5.9E-18*

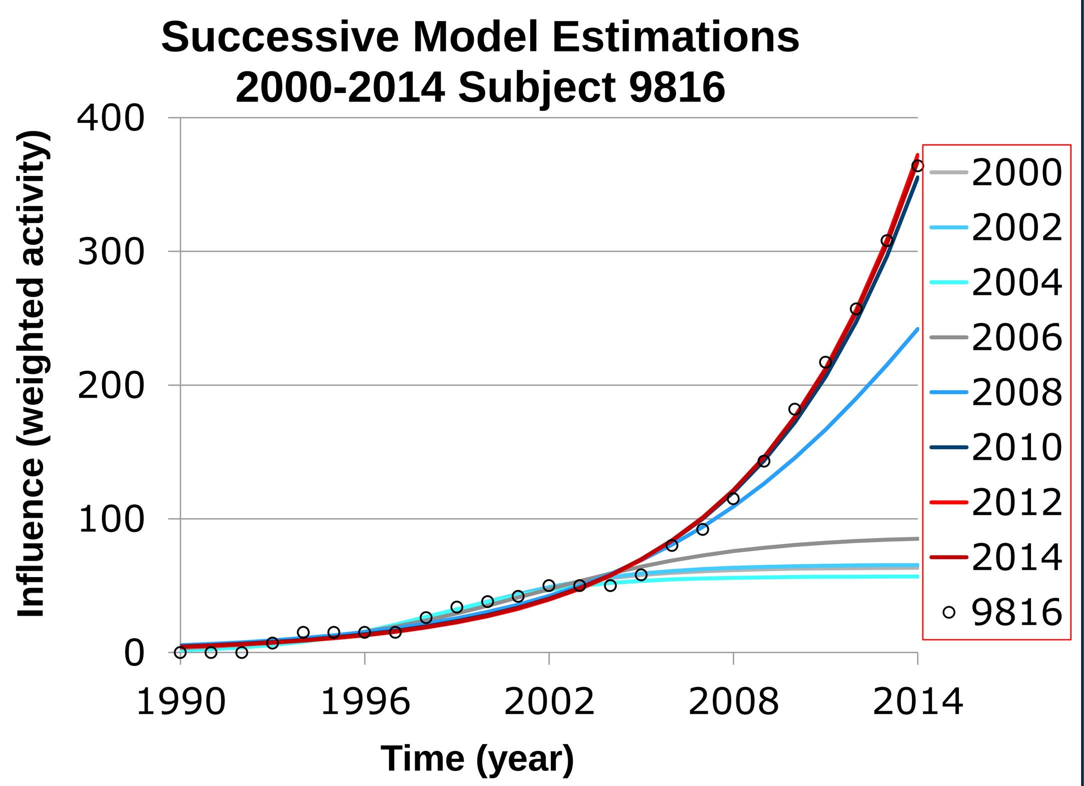
*Significance < 0.001

Results - Parameter Estimates

Subject	θ_1	θ_2	θ_3	t_{max}	Subject	θ_1	θ_2	θ_3	t_{max}
981	459	-14.4	0.49	2009	9844	586	-6.73	0.2	2012
982	547	-5.67	0.14	2020	9845	711	-5.63	0.13	2021
983	73	-16.3	0.53	2009	9846	59295	-14.4	0.24	2038
984	410	-13.7	0.42	2012	9848	45410	-10.9	0.15	2051
985	279	-16.7	0.53	2010	9857	579	-3.15	0.12	2006
986	454	-10.7	0.35	2009	9858	173	-9.89	0.31	2011
987	866	-9.94	0.26	2017	9861	197	-5.66	0.19	2009
9810	467	-6.45	0.23	2007	9862	180	-6.91	0.25	2007
9812	154	-7.47	0.29	2005	9864	1494	-9.26	0.25	2016
9813	1176	-4.73	0.16	2009	9893	170	-9.28	0.31	2009
9814	22790	-17.9	0.36	2029	9895	442	-12.4	0.35	2015
9816	21877	-10.6	0.19	2036	9896	697	-8.43	0.26	2012
9817	199	-6.25	0.2	2011	9899	436	-7.54	0.25	2010
9818	423	-15	0.44	2013	98111	360	-7.17	0.23	2010
9820	86218	-13.9	0.2	2048	98117	183	-8.8	0.28	2011
9821	729	-3.97	0.13	2010	98119	220	-4.97	0.2	2004
9826	340	-5.45	0.2	2006	98123	187	-11.3	0.38	2009
9827	208	-6.69	0.26	2005	98147	72	-7.69	0.26	2009
9828	633	-11.6	0.33	2014	98150	158	-14.9	0.48	2010
9829	179	-16.6	0.59	2007	98163	208	-9.95	0.37	2006
9833	700	-9.52	0.23	2021	98170	457	-4.92	0.17	2008
9837	296	-9.18	0.28	2011	98182	49495	-13.9	0.19	2054
9839	40	-14.8	0.48	2010	98205	37	-19.5	0.62	2010
9840	79	-11.3	0.43	2005	98218	228	-13	0.38	2013
9842	8715	-14.9	0.26	2036	98252	326	-12.3	0.42	2009



Sample Emerging KOL Prediction



KOL Emergence detected in the 2008 data and beyond

Model parameters θ_1 and t_{max} clearly indicate emergence starting in year 2008

Data Cutoff	θ_1	t_{max}	Yrs Data
2000	63	1999	8
2002	65	1999	10
2004	56	1998	12
2006	87	2001	14
2008	690	2018	16
2010	202,771	2049	18
2012	100,560	2044	20
2014	88,255	2044	22

Conclusion

This model provides a straightforward method of describing the influence growth profile of a HCP over time (relative to traditional field-based methods currently in use today). Using this model allows industry to view KOLs in terms of where they are in their influence life-cycle and estimate their maximum future productivity, allowing a life-sciences company to focus scarce and valuable resources on a set of target HCPs that should be more appropriate now *and* into the future.

Limitations

The model is purposely sensitive to the "tipping point" where an emerging KOL begins to appear. Currently, it does not limit t_{max} estimates. Further extensions to the model will be necessary to include this constraint.

References

1. Kroes M, C. O. (2011). Beyond KOL management: communities of practice as a new perspective on pharmaceutical market penetration. *Journal of Medical Marketing: Device, Diagnostic and Pharmaceutical Marketing*, Vol. 11, 1, 71–83.
2. BD., S. (2009 9, 4). An Exploratory Study of Key Opinion Leadership Management Trends among European Pharmaceutical Companies. *Journal of Medical Marketing: Device, Diagnostic and Pharmaceutical Marketing*, 291–300.
3. Farley K, H. A. (2014, 14:38). Gathering opinion leader data for a tailored implementation intervention in secondary healthcare: a randomised trial. *BMC Medical Research Methodology*, 1-7.
4. Verhulst, P. "Recherches mathématiques sur la loi d'accroissement de la population" [Mathematical Researches into the Law of Population Growth Increase]. *Nouveaux Mémoires de l'Académie Royale des Sciences et Belles-Lettres de Bruxelles* 18:1-42

Special Acknowledgements: Juan Nadal, Tara Pummer, Michael Toscani