Identification of Emerging Key Opinion Leaders Using Bibliometric Data Analysis

Authors: Gruber A.1, Taylor III H.2, Stapleton A.1, Booth K.1, Mello Jr. A.1, Rametta M.1
1Bayer HealthCare, Whippany, N.J., USA
2Bayer CropScience, Whippany, N.J., USA
3SteepRock Inc., Washington, Ch.T., USA

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

Methods - Model

1. **Fitting the Data:**
   - The model requires MySQL® "R" and Excel were used for model fitting.

2. **Logistic Growth Model:**
   - Functional Form: \( f(t) = \frac{1}{1 + \exp(-\theta t)} \)
   - Where: \( \theta \) is the average time the subject over time, \( \theta_1 \) is maximum influence attainable over the KOL, \( \theta_2 \) is an offset that determines the turning point of the curve in time.
   - \( \theta_1 \) controls the sensitivity of the curve to time.

3. **Bibliometric Data:**
   - 1990-2015
   - MySQL "R" public data

4. **Weighted KOL activity over time:**
   - 1970-2015
   - Initial growth, sociometric from currently to regulations of the market.

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

Results - Model Fit

- Model characteristics:
  - Well-known and understood (first published by Velhuis in 1984)²
  - Continuous, Non-Linear
  - Function in it
  - When parameterized it:
    - Initially grows gradually
    - Reaches a peak
    - Tapers off as time passes

- Data vs. Model Predictions
  - 2 Samples with Coefficients

- Model features:
  - Maximum influence is \( \theta_1 \)
  - Time frame of maximum productivity, \( t_{\text{max}} \), can be derived analytically.

Model of Influence Life-Cycle

Results - Parameter Estimates

- Subject: \( f(t) = \frac{1}{1 + \exp(-\theta t)} \), \( \theta_1 \), \( \theta_2 \), \( t_{\text{max}} \)

- Subject: \( f(t) = \frac{1}{1 + \exp(-\theta t)} \), \( \theta_1 \), \( \theta_2 \), \( t_{\text{max}} \)

- Data from 1990-2015

- General logistic growth model:

- Fitting the data:
  - \( \theta \) = 1, 1990-2015
  - \( \theta_1 \) = 1990-2015

- \( \theta_1 \) is the offset that determines the turning point of the curve in time.

- \( \theta_2 \) controls the sensitivity of the curve to time.

- Model fit:
  - \( \theta \) = 0.02, 1990-2015
  - \( \theta_1 \) = 0.02, 1990-2015

- Prediction of \( t_{\text{max}} \):
  - \( t_{\text{max}} \) = 2015

- Model parameters \( \theta_1 \), \( \theta_2 \), \( t_{\text{max}} \) clearly indicate emergence starting in year 2008.

- Sample emerging KOL prediction

- \( t_{\text{max}} \) detection in the data and beyond.

Methods - Data Set

- Wet Age-related Macular Degeneration (WAMD) Data
- Degeneration (WAMD) Database
- 50 KOLS: 1970-2014
- 50 Total of 2,872 data points

Conclusions
This model provides a straightforward method of describing the influence growth profile of a KOL over time relative to traditional field-based methods currently in use today. Using this model allows industry to view KCOLs in terms of where they are in their influence life-cycle and estimate their maximum productivity, allowing a life-science company to plan the efficient 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 \( t_{\text{max}} \) estimates. Further extensions to the model will be necessary to include this constraint.

References

Special Acknowledgments: Juan Nardé, Tara Furrer, Michael Toccas