Title

Life priorities in the HIV-positive Asians: a text-mining analysis in young vs. old generation

Authors

Chen, Wei-Ti
Barbour, Russell

Life priorities in the HIV-positive Asians: a text-mining analysis in young vs. old generation

Wei-Ti Chen & Russell Barbour

Routledge
Taylor & Francis Group

Life priorities in the HIV-positive Asians: a text-mining analysis in young vs. old generation

Wei-Ti Chen[a] and Russell Barbour[b]

[a]School of Nursing, Yale University, Orange, CT, USA; [b]Center for Interdisciplinary Research on AIDS (CIRA), Yale School of Public Health, New Haven, CT, USA

ABSTRACT

HIV/AIDS is one of the most urgent and challenging public health issues, especially since it is now considered a chronic disease. In this project, we used text mining techniques to extract meaningful words and word patterns from 45 transcribed in-depth interviews of people living with HIV/AIDS (PLWHA) conducted in Taipei, Beijing, Shanghai, and San Francisco from 2006 to 2013. Text mining analysis can predict whether an emerging field will become a long-lasting source of academic interest or whether it is simply a passing source of interest that will soon disappear. The data were analyzed by age group (45 and older vs. 44 and younger). The highest ranking fragments in the order of frequency were: "care", "daughter", "disease", "family", "HIV", "hospital", "husband", "medicines", "money", "people", "son", "tell/disclosure", "thought", "want", and "years". Participants in the 44-year-old and younger group were focused mainly on disease disclosure, their families, and their financial condition. In older PLWHA, social supports were one of the main concerns. In this study, we learned that different age groups perceive the disease differently. Therefore, when designing intervention, researchers should consider to tailor an intervention to a specific population and to help PLWHA achieve a better quality of life. Promoting self-management can be an effective strategy for every encounter with HIV-positive individuals.

HIV/AIDS is one of the most urgent and challenging public health issues in global health (World Health Organization, 2015). Globally, Asians HIV diagnosis slowly increased in recently years (WHO, 2015), however, specific research on Asians with HIV infections is sparse. Currently, the available data suggest that concerns for Asians living with HIV include acculturation stress, stigma, serostatus disclosure, medication access, and continuation of family obligations (Jones et al., 2010). Additional distresses include managing HIV-induced symptoms, antiretroviral therapy (ART) adherence and side effects of medication (Chen, Hoy, & Lewin, 2007).

As the CDC reported that 86% of the HIV-positive Asian and Pacific Islander (API) men were most likely transmitted via men who have sex with men (MSM) (Centers for Disease Control and Prevention, 2014), some of these HIV-positive API MSM were likely had married and have children (Chow, Koo, & Zhang, 2013). Given the pivotal position of family in Asian society and the marked physical, emotional, and social challenges faced by people living with HIV/AIDS (PLWHA), it is critical that we understand their life priorities and how these priorities impact their self-management strategies.

In this study, several in-depth interviews conducted among Chinese, Taiwanese, and Asians in San Francisco were combined to form a large qualitative dataset (Chen, 2013; Chen et al., 2013, 2015). Although Asians reflect diverse backgrounds, many of them follow similar cultural norms, especially with strong family values (Asian & Pacific Islander Wellness Center, 2011). Therefore, we hypothesized that older HIV-positive Asians will prioritize their needs around family themes.

We used text mining skill to explore the patterns of age differences among Asian PLWHA. The data were analyzed by two age groups (45 and older vs. 44 and younger). Text mining creates a data-driven, exploratory process of knowledge discovery where the focus is on finding and extracting useful patterns of information from large, complex databases (Hastie, Tibshirani, & Friedman, 2013).

Methods

Subjects

The study included 45 HIV-positive Asians (12 men and 33 women) who were diagnosed with HIV. The inclusion

criteria for the study were (a) 18 years or older, (b) diagnosed with HIV, (c) no cognitive impairment, and (d) willing to participate in an hour-long in-depth interview.

Procedure

The project was approved by the human subjects review board of the involved institutions. In-depth interviews were conducted by experienced researchers. Interviewers used a checklist during interviews to inquire about the participants' perceptions of their lives after their HIV diagnosis and what are the important things in their lives after diagnosis.

There were several recruitment sites including Beijing ($n = 10$), Shanghai ($n = 16$), Taipei ($n = 3$), and San Francisco ($n = 16$). Participants were approached through contacts at the local HIV infectious disease hospitals and community service centers where provide HIV care. Snowball sampling methods (Marshall, 1996) were used to identify potential study participants. PLWHA who were interested in the study met with the researchers, who then explained the study, answered questions, and obtained written consent.

All interviews were conducted in Mandarin or English, as the study participants preferred, and audio-recorded. Interviews conducted in English were transcribed verbatim. Interviews conducted in Chinese were transcribed into Mandarin verbatim then translated to English and re-transcribed. Interviews took one hour and were conducted in a private setting.

Analysis

Text mining is used to extract specific information from a large body of textual data (Castellani & Castellani, 2003; Reardon, 2014). We used statistical software known as the "*tm* package in *R*" (Feinerer, Hornik, & Meyer, 2008) for data analysis and categorized the contents of each interview by two age groups. The mean age was 35.46 in the 44-and-younger group, and 53.14

years old in the 45-and-older group. More demographic data was presented at Table 1.

In the first process, the words in each sentence were separated. Words that had the same meaning were counted as the same word; for example, "HIV" and "AIDS" were both counted as "HIV", while "medicine" or "treatment" or "therapy" were counted as "medicine". Next, articles (a, an, the), punctuation marks, and numbers were deleted, leaving only meaningful words. The meaningful words remaining after the segregation process are called "fragments". In this case, text mining elicited 113 fragments, which were subjected to correspondence analysis for chosen effective characteristics.

Then, we conducted dissimilarity analysis on each case based on age groups to determine each participant's family support. Dissimilarity analysis is defined so as to express the similarity (pair-to-pair) between the sets (Feinerer & Hornik, 2014). In similarity/dissimilarity analysis, participant responses of different groups were compared across a single set of questions and were measured using the extended Jaccard coefficient, as implemented in the *tm* package for *R* statistical software (Feinerer & Hornik, 2014). The Jaccard coefficient measures similarity/dissimilarity as the intersection divided by the union of the objects. When applied to text documents, this metric compares the sum weight of shared terms to the sum weight of terms that present in either of the two documents but are not shared (Hastie et al., 2013).

Results

After doing correspondence analysis between fragments and age groups, we selected those fragments that represented concerns participants had regarding their lives. After removed the sparse terms, the software

Table 1. Description of the study participants, by sites.

Characteristic	San Francisco, US (N = 16)	Shanghai, China (N = 16)	Beijing, China (N = 10)	Taipei, Taiwan (N = 3)
Age range (years old)	36–69	30–58	25–59	44–65
45-year-old and above	10	8	1	2
44-year-old and younger	6	9	8	1
Marital status				
Married	0	7	5	1
Not married	16	9	5	2

Note: Countries of origin were Taiwan (3), Philippines (7), Vietnam (4), Indonesia (1), Japan (1), Hawaii (native) (1), Malaysia (1), Cambodia (1), and China (26).

Table 2. Fragments selected by age.

Order	Total participants	44 years and younger	45 years and above
		Significant fragments	
1	Care	People asked	Care
2	Daughter	Daughter	Disease
3	Disease	Doctors	Family
4	Family	Family	Feel
5	Feel	Feel	HIV
6	HIV	HIV	Hospital
7	Hospital	Hospital	Know
8	Husband	Husband	Me
9	Medicine	Know	People said
10	Money	Medicines	Son
11	Son	Patients	Tell/disclosure
12	Tell/disclosure	People said	Thought
13	Thought	Thought	Told
14	Want	Told	Want
15	Years	Want	Years

presented the result for these 45 cases. In this analysis, we kept just the 1% most common words from the study cases. A process was then run to choose those significant (effective) words or phrases that showed up in the text at least 281 times. The highest and lowest ranking words in a category were then chosen (see Table 2). The word frequencies were reduced in proportion to the number of documents that fell into the classifications under consideration.

We next tested the process on the most frequent terms (terms that were present in the text more than 180 times) within 2 groups. Table 2 presented in the highest and lowest ranking in the category.

Discussion

Our analysis showed that there were differences among the set of words for younger vs. older PLWHA. Participants in the younger group were focused mainly on disease disclosure, their families, and their financial condition. Many phrases regarding children demonstrated concerns about the future. As is common with chronic diseases, parents worried that they could no longer perform their role of caring for the family (Williams, Skirton, Barnette, & Paulsen, 2012). Asian PLWHA in both groups concerned on how the disease will impact with their families. They often struggle with the problem of whether to disclose their HIV status or to simply try to keep living with HIV without disclosing (Li, Lin, Wu, Lord, & Wu, 2008).

How to disclose is another issue. PLWHA often wonder whether they should disclose partially (e.g., disclose that they had some unspecified blood-related disease) or disclose fully (mention HIV and AIDS) (Simoni et al., 2015). Under these conditions, family and social support become key regimens to be used in preventing the situation from getting worse. In older PLWHA, social support was one of the main concerns (Ball, 2014). Asian elders are expected to stay close to younger family members, especially their sons (Laidlaw, Wang, Coelho, & Power, 2010). Yet an older person with HIV might be afraid that he might transmit the disease to loved ones. Therefore, social support should be evaluated during each visit (Dury, 2014).

Younger PLWHA tend to be more focused on their current living conditions, which are impacted by socioeconomic status, severity of condition, and family stability (Ladefoged, Andersson, Koch, Rendal, & Rydbacken, 2012). In addition, fatigue, negative feelings, lowered cognitive ability, decreased financial resources, and ongoing HIV symptoms also contribute to diminished quality of life (QOL) (Skevington, 2012). Studies have shown that social support and personal relationships are important in helping PLWHA to find continued meaning in life (Margalho, Pereira, Ouakinin, & Canavarro, 2011).

Limitations

There were several limitations in this study. First, the study was conducted at several sites, where participants might have had differing experiences and bias with HIV experiences. Thus, it is difficult to generalize the results of this study to a larger population. However, they were all holding strong family value and needing similar family support. Second, there were more women in this project, which might skew the analysis. Finally, the statistical software "R", which currently accepts only English-language input. Because of this English-only restriction, several research assistants had to be recruited to translate the Chinese in-depth transcriptions to English. These translators had different skill levels and translating styles. For example, one might have translated a Chinese-language comment as "was afraid" while another would use "had fear", one might have used the term "provider" while another used "doctor". We tried to combine and reconcile all similarly expressed ideas during the data analysis, but there were some instances where we were not able to do that and some other cases where we might have overlooked.

Conclusion

Data mining give researchers a new lens to view the life priorities of PLWHA. Using this lens, researchers can compare these priorities as they get older. Based on the analysis, we conclude that a disclosure intervention can enhance family support. Also, a coordinated approach involving all healthcare providers should be considered when providing interventions to support the socially isolated PLWHA (Dury, 2014). This life priority factors had shown a new approach to integrate follow-up counseling with disclosure and social support intervention in HIV-positive Asian population.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

References

Asian & Pacific Islander Wellness Center. (2011). *HIV and Asian & Pacific Islanders: The epidemic in San Francisco.* Retrieved July 5, 2013, from http://www.apiwellness.org/extras/2011/sf_factsheet_REV2011.pdf

Ball, S. C. (2014). Increased longevity in HIV: Caring for older HIV-infected adults. *Care Management Journals*, 15(2), 76–82.

Castellani, B., & Castellani, J. (2003). Data mining: Qualitative analysis with health informatics data. *Qualitative Health Research*, 13(7), 1005–1018.

Centers for Disease Control and Prevention. (2014). *HIV infection among Asians in the United States and dependent areas.* Atlanta, GA: Centers for Disease Control and Prevention. Retrieved from http://www.cdc.gov/hiv/risk/racialethnic/asians/

Chen, L. F., Hoy, J., & Lewin, S. R. (2007). Ten years of highly active antiretroviral therapy for HIV infection. *Medical Journal of Australia*, 186(3), 146–151.

Chen, W.-T. (2013). Chinese female immigrants English-speaking ability and breast and cervical cancer early detection practices in the New York metropolitan area. *Asian Pacific Journal of Cancer Prevention*, 14(2), 733–738.

Chen, W. T., Guthrie, B., Shiu, C. S., Wang, L., Weng, Z., Li, C. S., … Luu, B. V. (2015). Revising the American dream: How Asian immigrants adjust after an HIV diagnosis. *Journal of Advanced Nursing.* doi:10.1111/jan.12645

Chen, W. T., Lee, S. Y., Shiu, C. S., Simoni, J. M., Pan, C., Bao, M., & Lu, H. (2013). Fatigue and sleep disturbance in HIV-positive women: A qualitative and biomedical approach. *Journal of Clinical Nursing*, 22(9–10), 1262–1269. doi:10.1111/jocn.12012

Chow, E. P., Koo, F. K., & Zhang, L. (2013). Are wives of gay men becoming the next target of HIV infection in China? *Sexually Transmitted Diseases*, 40(12), 964–965. doi:10.1097/OLQ.0000000000000050

Dury, R. (2014). Social isolation and loneliness in the elderly: An exploration of some of the issues. *British Journal of Community Nursing*, 19(3), 125–128. doi:10.12968/bjcn.2014.19.3.125

Feinerer, I., & Hornik, K. (2014). *Text mining package: R package version 0.5-10.* Retrieved August 17, 2015, from http://CRAN.R-project.org/package=tm

Feinerer, I., Hornik, K., & Meyer, D. (2008). Text mining infrastructure in R. *Journal of Statistical Software*, 25(5), 1–54.

Hastie, T., Tibshirani, R., & Friedman, J. (2013). Boosting and additive trees. In T. Hastie (Ed.), *The elements of statistical learning: Data mining, inference, and prediction* (pp. 350–352). New York: Springer.

Jones, D. L., Owens, M. I., Lydston, D., Tobin, J. N., Brondolo, E., & Weiss, S. M. (2010). Self-efficacy and distress in women with AIDS: The SMART/EST women's project. *AIDS Care*, 1–10. doi:10.1080/09540121.2010.484454

Ladefoged, K., Andersson, M., Koch, A., Rendal, T., & Rydbacken, M. (2012). Living conditions, quality of life, adherence and treatment outcome in Greenlandic HIV patients. *International Journal of Circumpolar Health*, 71, 2851. doi:10.3402/ijch.v71i0.18639

Laidlaw, K., Wang, D., Coelho, C., & Power, M. (2010). Attitudes to ageing and expectations for filial piety across Chinese and British cultures: A pilot exploratory evaluation. *Aging & Mental Health*, 14(3), 283–292. doi:10.1080/13607860903483060

Li, L., Lin, C., Wu, Z., Lord, L., & Wu, S. (2008). To tell or not to tell: HIV disclosure to family members in China. *Developing World Bioethics*, 8(3), 235–241. doi:10.1111/j.1471-8847.2007.00220.x

Margalho, R., Pereira, M., Ouakinin, S., & Canavarro, M. C. (2011). Adherence to HAART, quality of life and psycho-pathological symptoms among HIV/AIDS infected patients. *Acta Medica Portuguesa*, 24(Suppl. 2), 539–548.

Marshall, M. N. (1996). Sampling for qualitative research. *Family Practice*, 13(6), 522–526.

Reardon, S. (2014). Text-mining offers clues to success. *Nature*, 509(7501), 410. doi:10.1038/509410a

Simoni, J. M., Yang, J. P., Shiu, C. S., Chen, W. T., Udell, W., Bao, M., … Lu, H. (2015). Nurse-delivered counselling intervention for parental HIV disclosure: Results from a pilot randomized controlled trial in China. *AIDS*, 29(Suppl. 1), S99–S107. doi:10.1097/QAD.0000000000000664

Skevington, S. M. (2012). Is quality of life poorer for older adults with HIV/AIDS? International evidence using the WHOQOL-HIV. *AIDS Care*, 24(10), 1219–1225. doi:10.1080/09540121.2012.661838

Williams, J. K., Skirton, H., Barnette, J. J., & Paulsen, J. S. (2012). Family carer personal concerns in Huntington disease. *Journal of Advanced Nursing*, 68(1), 137–146. doi:10.1111/j.1365-2648.2011.05727.x

World Health Organization. (2015). *HIV/AIDS.* Retrieved from HIV/AIDS website: http://www.who.int/topics/hiv_aids/en/