Using natural language processing to determine predictors of healthy diet and physical activity behavior change in ovarian cancer survivors

John Culnan,[1] Tracy E. Crane,[2,3,4] Rebecca Sharp,[1] Sarah J. Wright,[2,3] Ti-Tai Wang,[5] G. Hagan Franks,[6] Chris Klimowski,[6] Nirav Merchant,[6] Steven J. Bethard[5]

Author Affiliations: 1. Department of Linguistics, College of Social and Behavioral Sciences, University of Arizona 2. College of Nursing, University of Arizona 3. University of Arizona Cancer Center 4. Mel and Enid Zuckerman College of Public Health, University of Arizona 5. School of Information, College of Social and Behavioral Sciences, University of Arizona 6. Data Science Institute, University of Arizona

Introduction

- Approximately 50 percent of new cancer cases could be eliminated through a combination of healthy behaviors (e.g., physical activity and healthy diet)
- Cancer survivors are at high risk for developing new and recurrent cancer
- A significant percentage of cancer survivors are not attaining the cancer preventive guidelines of healthy diet and physical activity
- Telephone-based lifestyle interventions have demonstrated effectiveness in helping survivors meet cancer preventive guidelines however these trials are labor intensive and expensive to deliver, limiting their potential for broad dissemination
- Machine learning and Natural Language Processing are analytical techniques that automatically learn from direct and indirect patterns in data
- Machine learned algorithms may be used to analyze speech and aid in predicting who may be at risk of poor adoption of healthy lifestyle behaviors
- Vocalics, acoustic cues to emotion, personality, and attitudes in the speech signal, may be used to help distinguish individual levels of motivation and unspoken difficulties with reaching goals of a lifestyle intervention

Purpose

- To explore the use of speech technology and Natural Language Processing in evaluating language and vocalics as predictors of behavior change in ovarian cancer survivors participating in a lifestyle intervention

LIVES Study

- The Lifestyle Intervention for Ovarian Cancer Enhanced Survival (LIVES) study is testing whether women randomly assigned to a lifestyle intervention promoting a high vegetable, fruit and fiber and low-fat diet and increased physical activity will have increased progression-free survival as compared to women assigned to an attention control
- Daily intervention goals: 4 servings of vegetables, 2 servings of fruit, 30 grams of fiber, 20% of calories from fat and an additional 4,000 steps per day above baseline.
- Motivational interviewing, a directed, patient-centered coaching approach, is used to elicit behavior change for women assigned to the intervention
- Recorded coaching calls with study participants are administered on a sliding scale over two years; recordings scored (*a priori*) to assess coaching protocol fidelity

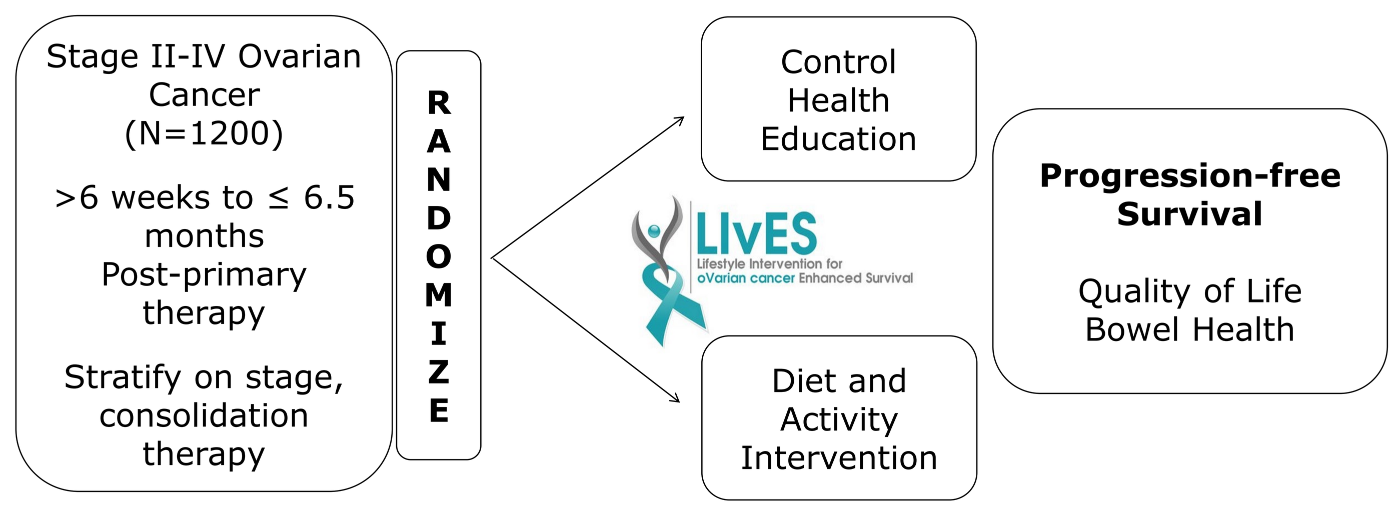

Figure 1: LIVES Study Schematic

Methods

- 246 randomly selected, recorded telephone coaching sessions from women participating in the LIVES study were used for this analysis; coaching sessions followed a coaching protocol with intervention fidelity scores set *a priori*

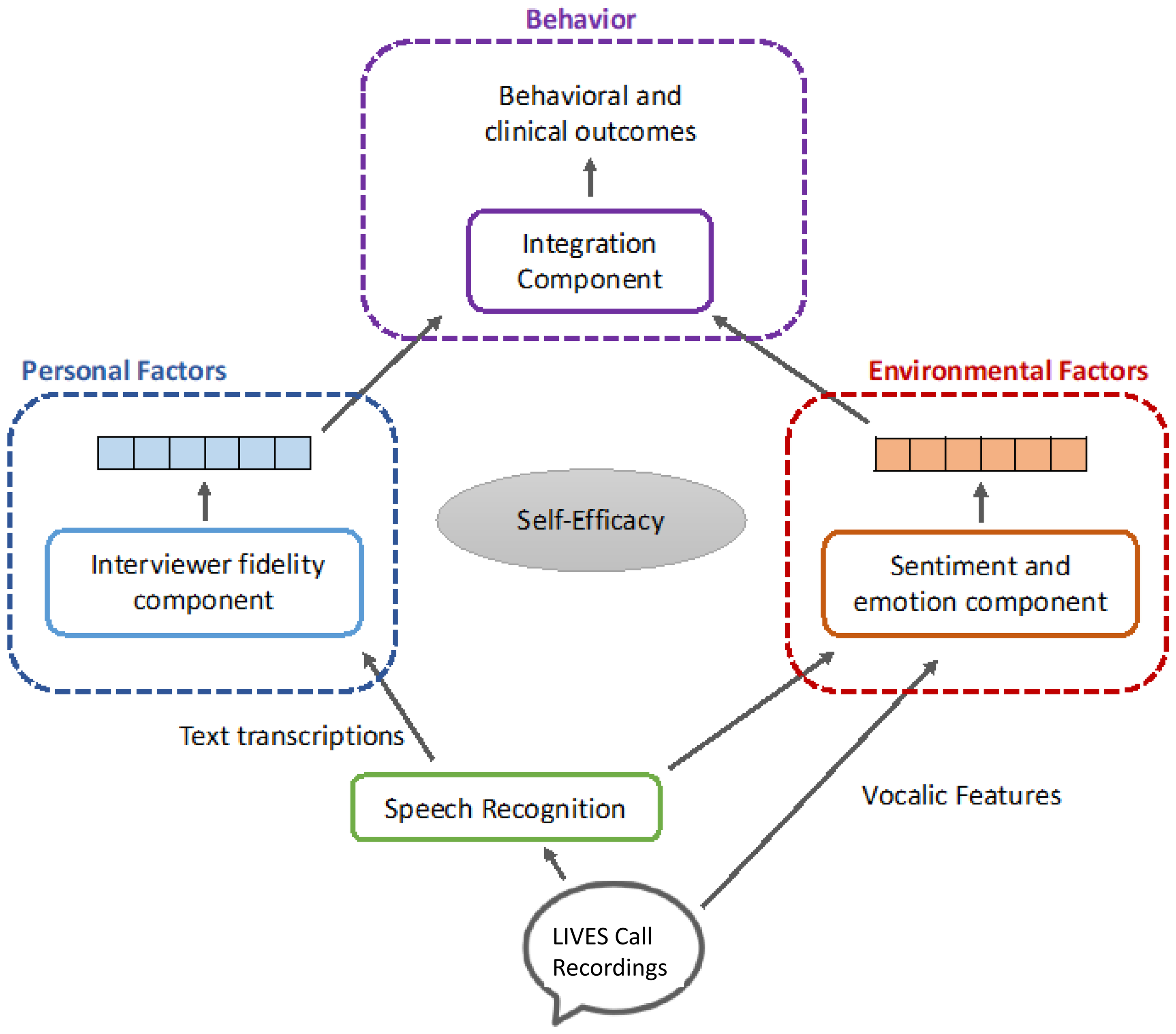

Figure 2: Integrated conceptual framework of Natural Language Processing and Machine Learning with Social Cognitive Theory for behavior change.

Speech-to-text analysis

- Three automated speech recognition programs, Google Cloud Text-to-Speech, AWS Transcribe, and Watson Speech to Text, were tested

Text modeling

- Logistic regression classifiers were trained to predict fidelity variables, using a bag-of-words representation of the transcriptions as input
- These models were compared to a baseline model that simply predicts the most frequent value

Speech modeling

- Using the OpenSMILE acoustic feature extraction library, the audio was analyzed for how well different aspects of the speech signal (e.g., pitch, spectral energy) reflect low vs. high participant achievement
- Multiple linear regression was performed to calculate the usefulness of numerous predictor variables as predictors of participants' intake of vegetables, fiber, and fat, as well as MET Hrs/week (a measure of physical activity) at the six-month follow-up assessment
- The model included measurements of diet intake and physical activity at the start of the program, demographic information, and measures of speech loudness, pitch, and voice quality

Acknowledgement of funding

The authors acknowledge the important role the women participating in the Lifestyle Intervention for oVarian Cancer Enhanced Survival [(LIvES)-GOG/NRG 0225] in providing data for this work. The LIvES-GOG/NRG 0225 program is funded by the National Institutes of Health National Cancer Institute (1R01 CA186700) and the Behavioral Measurement and Interventions Shared Resource at the University of Arizona Cancer Center Support Grant (P30 CA023074) and a University of Arizona Health Sciences Career Development award to Tracy Crane (1470602-CDA27)

Results

Speech-to-text analysis

- Google Cloud Text-to-Speech performed the best with 76% accuracy in transcription (WER), and 32% accuracy in speaker identification (DER)

Text modeling

- Text classifiers trained and tested (via leave-one-out cross-validation) on the 49 calls that were both transcribed and had fidelity scores achieved good predictive power over several fidelity variables

Fidelity score variable	Input portion of transcript	Classifier accuracy	Baseline accuracy
Review of FIBER goal	First half	0.71	0.51
Review of FAT goal	Full transcript	0.71	0.57
Numerical review of ANY goal - servings	First half	0.77	0.67
Numerical review of ANY goal - grams	Full transcript	0.81	0.65

Table 1: The performance of Bag-of-words classifier on four fidelity variables showing great improvement

Speech modeling

- Initial regression on 48 participants' check-ins demonstrated that more successful participants at the 6-month check-in:
  - More variable pitch,
  - Spoke more quietly, with less change in volume
  - Less variable average voice quality, with moments of large quality change

Goal	Acoustic cue	Effect	P-value
Change in fat consumption	Q3 loudness	Quieter=smaller decrease	p<.01
	Mean rate of change loudness	Faster rate of change=smaller decrease	p<.01
	Mean jitter	More variable voice quality=smaller decrease	p<.05
Change in fiber consumption	Q3 loudness	Quieter=greater increase	p<.01
	Q1 rate of pitch change	More variable pitch=greater increase	p<.02
Change in vegetable servings	Q3 loudness	Quieter=greater increase	p<.05
	Q3 pitch	Higher=greater increase	p<.001
	Q1 rate of pitch change	More variable=greater increase	p<.01

Table 2: Acoustic cues in participant speech that were predictive of changes in fat, fiber, and vegetable consumption by participants at the 6-month check-in

Conclusions

- Speech technology and Natural Language Processing hold high potential for identifying characteristics of language used in coaching calls that can be used to tailor and improve lifestyle behavior interventions for those who are at high risk for not meeting their goals

Future Directions

- Next steps will include analysis of the more than 33,000 recorded hours of LIVES calls for language and sentiment in relation to diet and physical activity behavior change