Effects of Aging on the Cortical Representation of Continuous Speech

I.M Dushyanthi Karunathilake[1], Jason L. Dunlap[2], Janani Perera[2], Alessandro Presacco[3], Lien Decruy[3], Samira Anderson[2], Stefanie E. Kuchinsky[4], Jonathan Z. Simon[1,3,5]

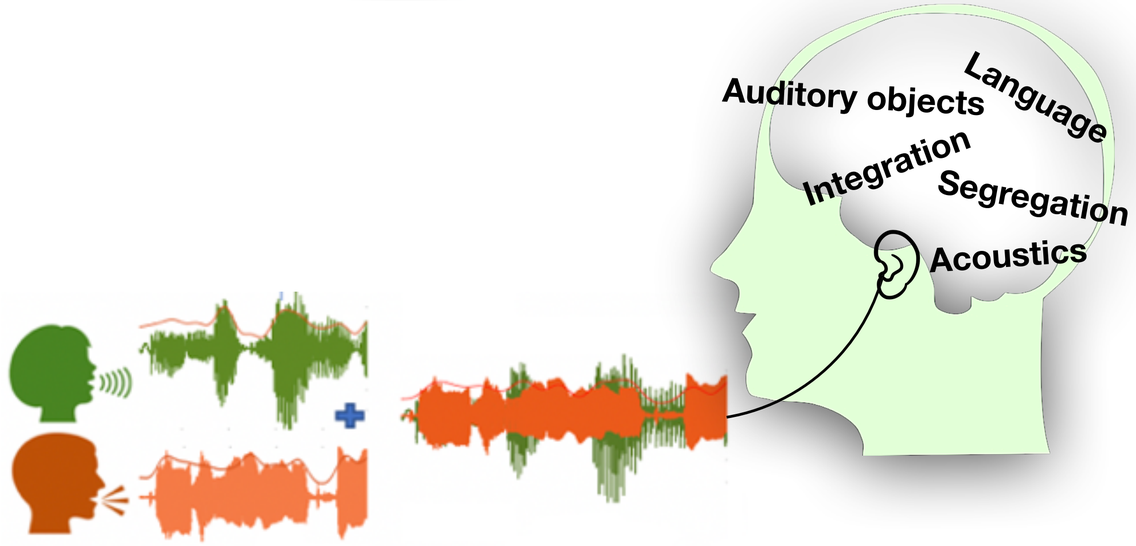
[1]Department of Electrical and Computer Engineering, [2]Department of Hearing and Speech Sciences, [3]Institute for Systems Research, [5]Department of Biology, University of Maryland, College Park, MD 20742, USA
[4]Audiology and Speech Pathology Center, Walter Reed National Military Medical Center, Bethesda, MD 20889, USA

Introduction

- Older adults often report difficulty in understanding speech in noise [1]
- Difficulties may arise from age-related physiological changes and temporal processing deficits
- Previous work has shown that older adults exhibit exaggerated cortical responses

Motivation

- At what latencies, age related processing differences occur?
- How does the task difficulty change the neural responses?
- How are the foreground (FG) and background (BG) speakers represented neurally?



1. Gordon-Salant, S., Fitzgibbons, P. J., & Yeni-Komshian, G. H. (2011). Auditory temporal processing and aging: Implications for speech understanding of older people. *Audiology Research*, *1*(1S), e4. https://doi.org/10.4081/audiores.2011.e4

Method

Participants: Native English speakers

- 18 younger adults (age: 17-26 y)
- 17 older adults (age: 65-78 y)
- Normal hearing (125-4000 Hz thresholds $\leq$ 25 dB HL)

Task: Listening to 1 minute long speech segments from an audio book

- Clean speech
- Mixed speech (Male speaker vs female speaker) [0 dB, -6 dB]
- Attend to male or female speaker ignoring the other

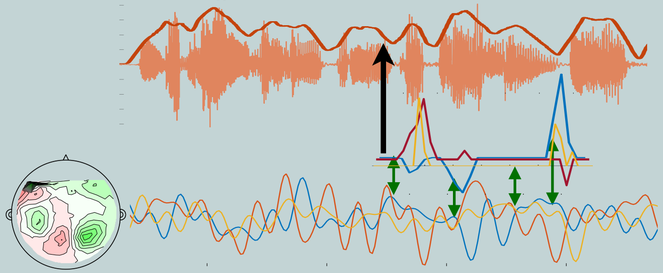
Data: MEG data

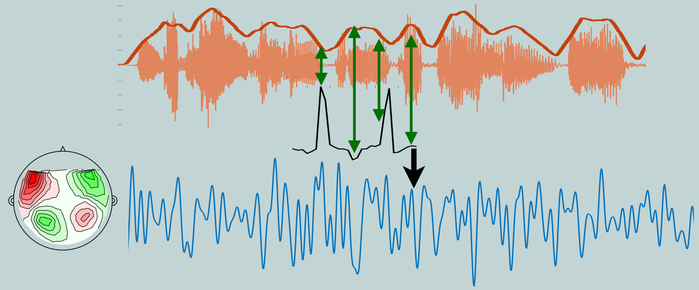
- Band pass filter 1-10 Hz

Preprocessing: Denoised by Time-Shift PCA (TSPCA), Sensor noise suppression (SNS) and Denoising Source Separation (DSS)

Cortical representation analysis

- Low frequency (1-10 Hz) log speech envelope
- Boosting algorithm with 5-fold cross validation
- 500ms integration window
- Statistical significance was evaluated by Linear mixed effect models





Stimulus Reconstruction

- First 6 DSS components filtered data for the reconstruction
- Attended and Unattended speaker envelopes reconstructed separately
- Reconstruction accuracy is estimated by the Pearson correlation coefficient between the reconstructed and the true speech envelope

Temporal Response Function (TRF) [2]

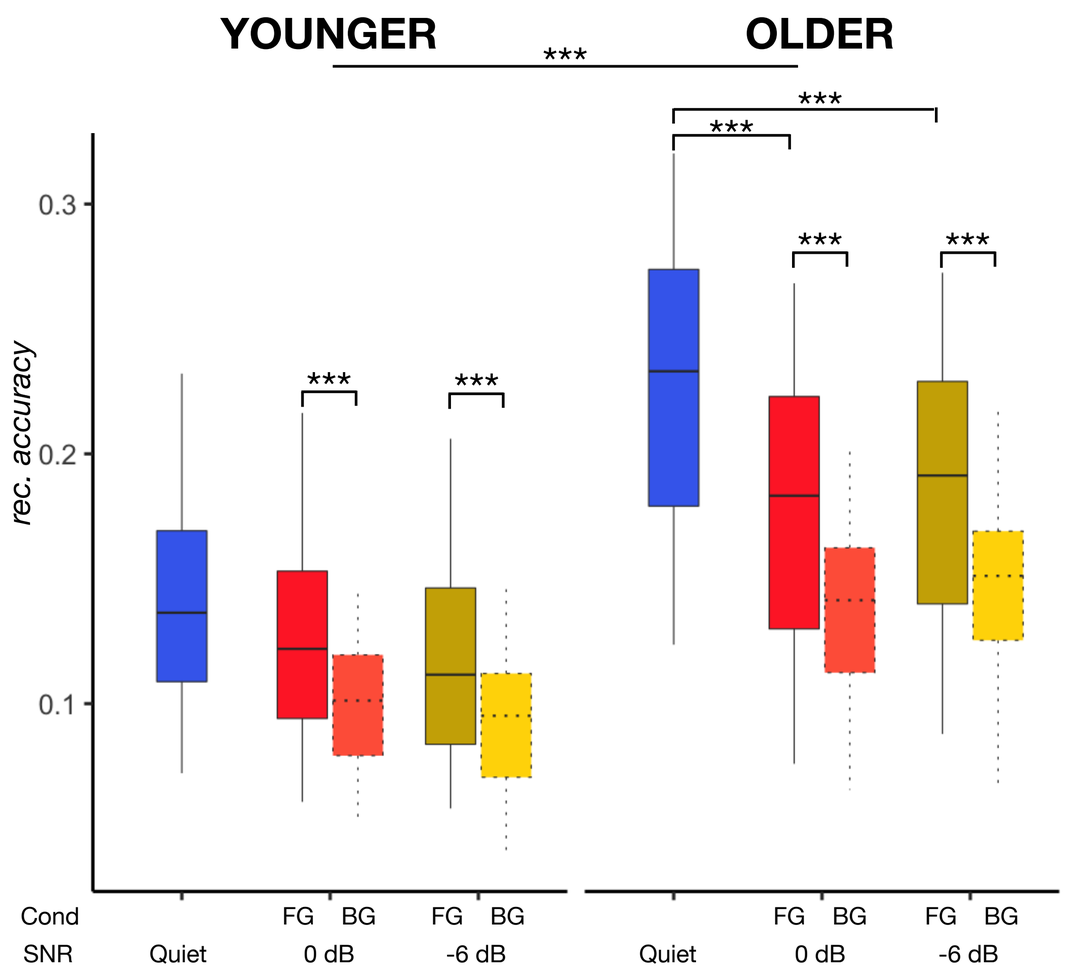
- 1st DSS component filtered neural data is used as the auditory response
- TRF is estimated as the linear filter that transforms the speech envelope to the neural response

$$DSS_1'(t) = \sum_{\tau} TRF_{FG}(\tau) Env_{FG}(t-\tau) + \sum_{\tau} TRF_{BG}(\tau) Env_{BG}(t-\tau) + \varepsilon(t)$$

- TRF has three prominent peaks
  - ~50 ms ; M_{50} (a positive peak)
  - ~100 ms; M_{100} (a negative peak)
  - ~200 ms; M_{200} (a positive peak)

2. Ding, N., & Simon, J. Z. (2012). Emergence of neural encoding of auditory objects while listening to competing speakers. *Proceedings of the National Academy of Sciences*, *109*(29), 11854–11859. https://doi.org/10.1073/pnas.1205381109

Stimulus Reconstruction



$$rec.accuracy \sim 1 + Age + SNR + Age * SNR + Cond + (1 + Cond \mid Subj)$$

*p<0.05, **p<0.01, ***p<0.001

Younger vs Older

Older reconstruction better than younger

- Confirms the previous results [3,4]

Possible Explanations:

- Age related changes e.g., excitation/ inhibition imbalance
- Recruitment of additional top-down resources
- Increased attention

Task Difficulty

Age*SNR is significant
Quiet reconstruction better than 0 dB/-6 dB only in older adults

Related finding:

- Background noise significantly worsens speech intelligibility in older listeners

Foreground vs Background

Foreground reconstruction better than Background for both groups

Possible Explanations:

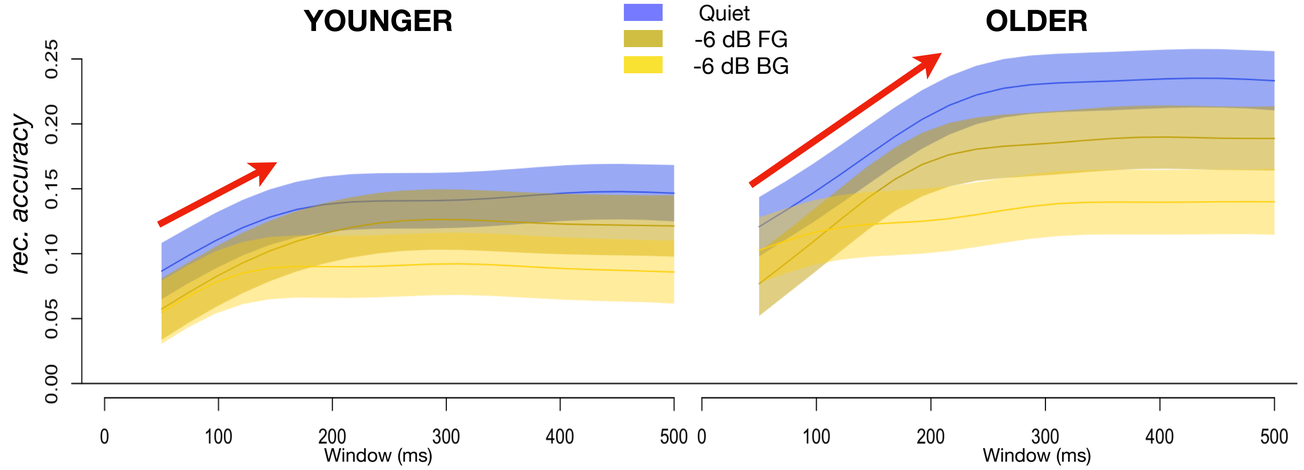
- Selective attention
- Separation into distinct sources

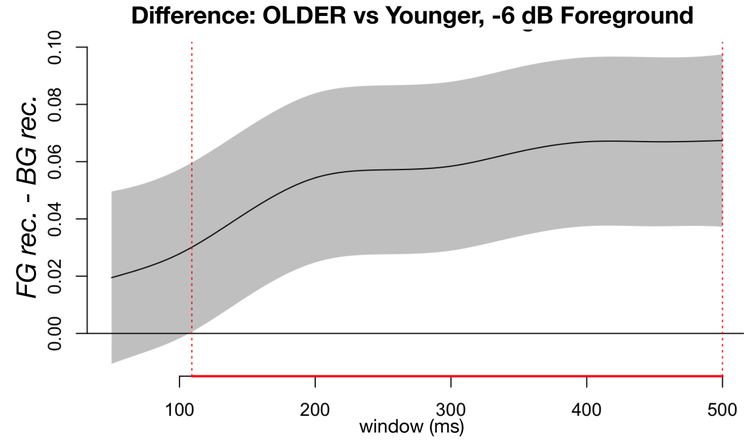
3. Presacco, A., Simon, J. Z., & Anderson, S. (2016). Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. *Journal of Neurophysiology*, *116*(5), 2346–2355. https://doi.org/10.1152/jn.00372.2016
4. Decruy, L., Vanthornhout, J., & Francart, T. (2019). Evidence for enhanced neural tracking of the speech envelope underlying age-related speech-in-noise difficulties. *Journal of Neurophysiology*, *122*(2), 601–615. https://doi.org/10.1152/jn.00687.2018

Stimulus Reconstruction - Integration window analysis

- Method:- Speech envelope reconstruction using 50-500 ms integration windows
- Statistical Analysis:- Generalized additive Mixed models (GAMMs) time series

$$Rec \sim \beta_0(Age * SNR * Attn) + s(Window, Age * SNR * Attn) + (s(Window) \mid Subj)) + (Window \mid Item) + \epsilon$$





Younger vs Older

Reconstruction takes more time for older adults

Possible Explanations:
- Recruitment of additional top-down resources
- Attentional gain

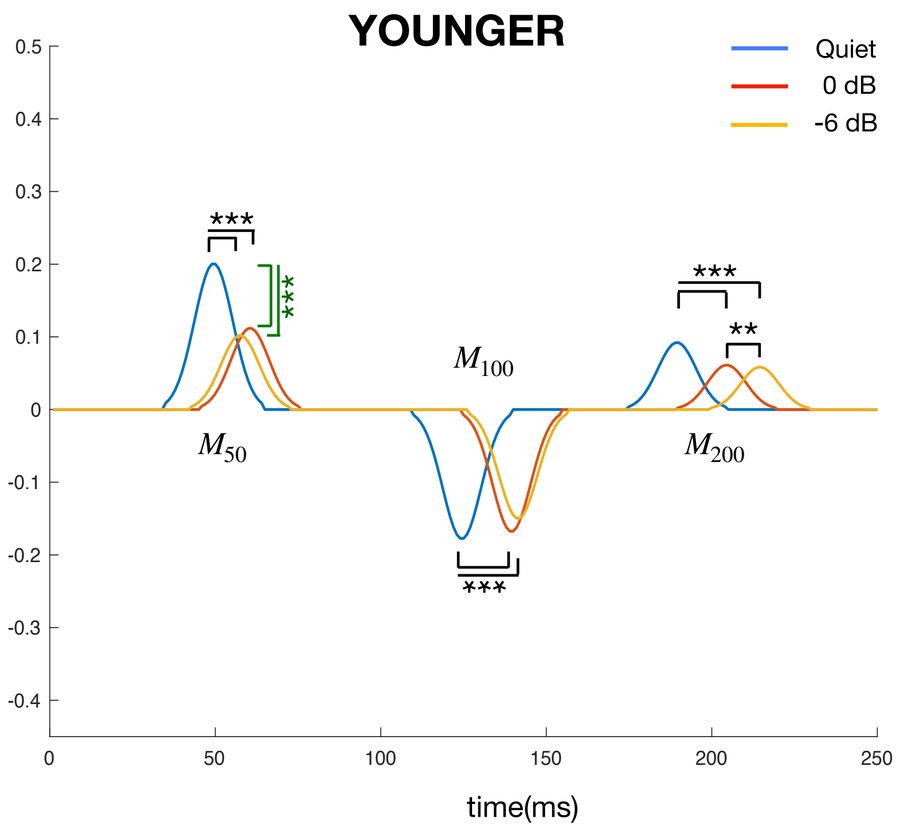
Over representation starts as early as ~100ms in older adults

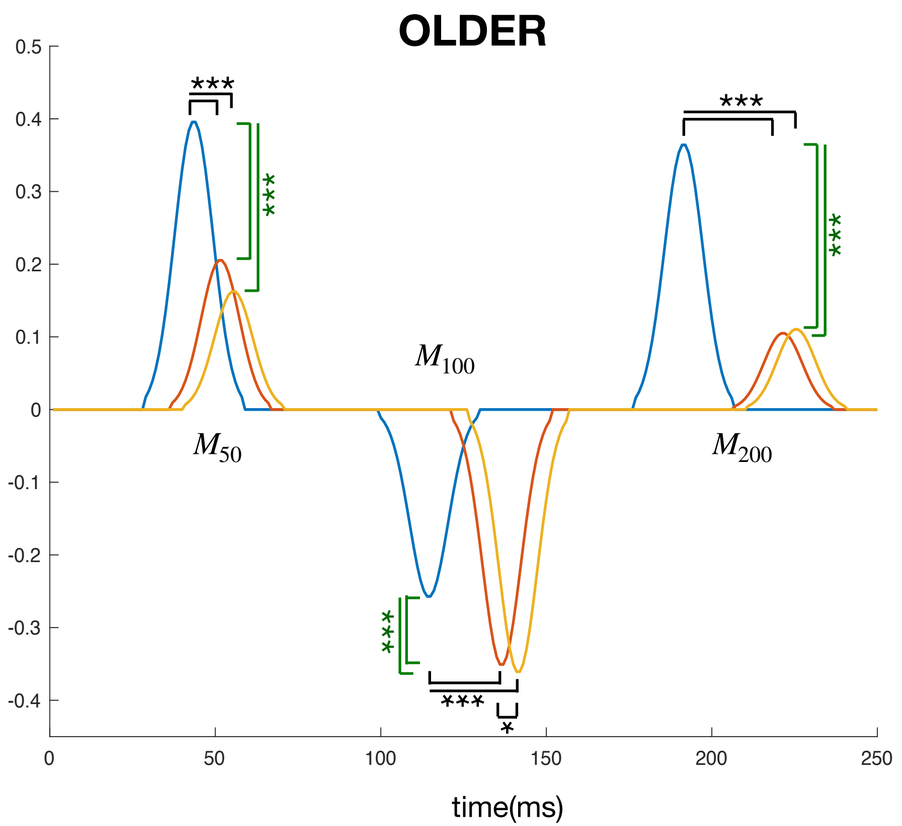
Possible Explanations:
- Excitation and inhibition imbalance

Ceiling: Younger (100-200 ms) < Older (200-300 ms)

Possible Explanations:
- Additional processing ~200ms in older adults to compensate for the temporal processing deficits

Temporal Response Function (TRF) - Foreground





Plots contain illustrative TRFs

**p<0.05, **p<0.01, ***p<0.001*

Younger vs Older

Older amplitudes bigger than younger amplitudes

Possible Explanations:

- Age related changes e.g., excitation/ inhibition imbalance

M200 peak difference

Possible Explanations:

- More resources being used for processing

Task Difficulty

M100 increases only in older adults

Possible Explanations:

- Greater attention
- Listening asymmetry

M200 amplitude decreases

Possible Explanations:

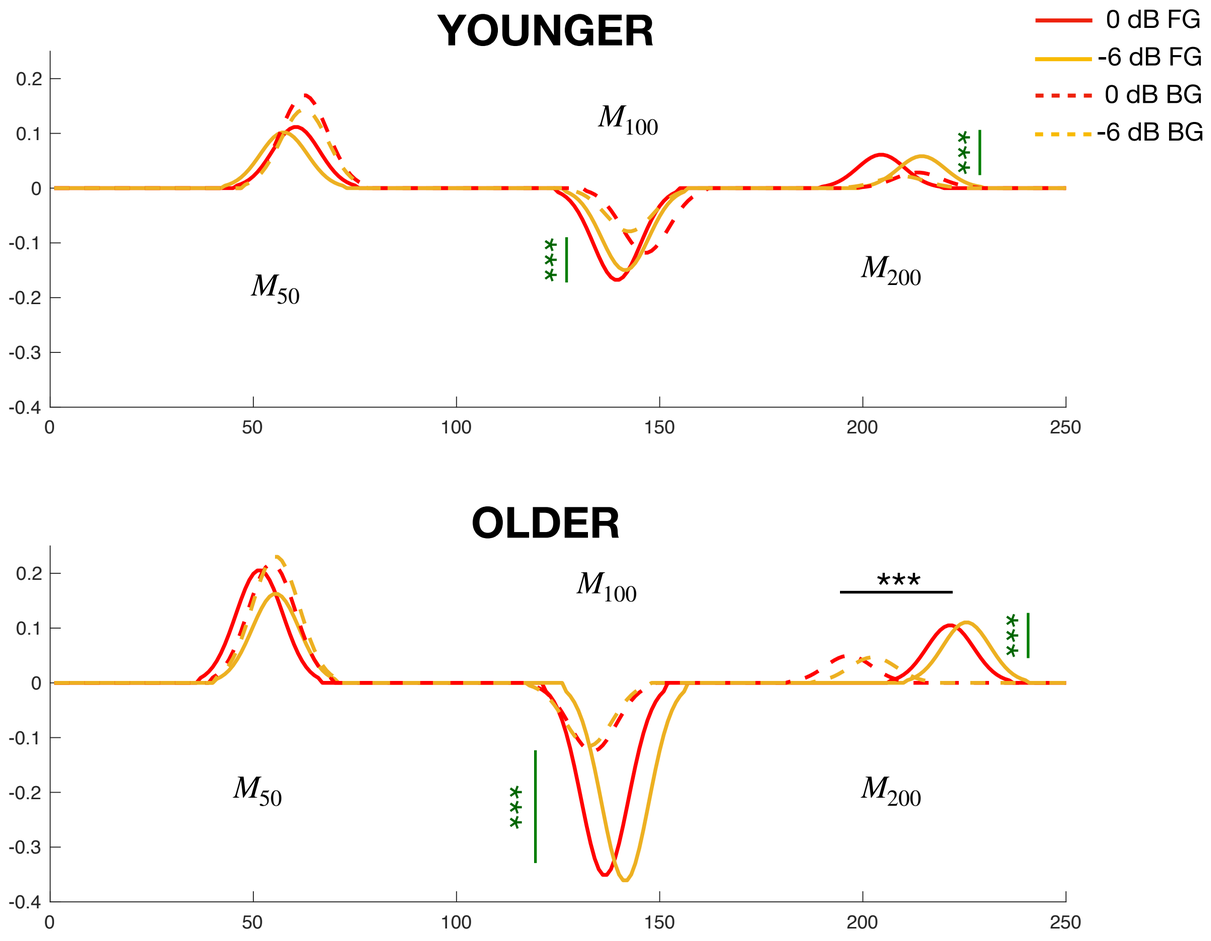
- Modulated by late neural mechanisms

Quiet peaks early than 0 dB/-6 dB peaks

Possible Explanations:

- Harder the task, takes more time to process

Temporal Response Function: Foreground Vs Background



Plots contain illustrative TRFs

**p<0.05, **p<0.01, ***p<0.001*

Younger vs Older

No difference in peaks

Foreground vs Background

No significant difference between M50 for both groups

Possible Explanations:

- Early response is shared between Foreground and Background

Foreground > Background for both M100 and M200 and both groups

Possible Explanations:

- Modulated by attention

Foreground M250 latency> Background M250 latency only in older adults

Possible Explanations:

- Foreground is processed for longer time by recruiting additional resources

Conclusion

- Older adults' neural response robustly tracks the speech envelope, and to a greater extent than younger adults, possibly due to several mechanisms
- M200 peak is late enough to be modulated by many compensatory mechanisms
- Early activity, i.e., the M50, is not modulated by attention, while late activity, M100 and M200, is
- More difficult tasks produce longer latencies
- Altogether, despite impaired speech intelligibility in noise, time locked speech responses are exaggerated in older adults compared to younger

Acknowledgement

This work was supported by the National Institutes of Health grants P01-AG055365 and R01-DC014085

The views expressed in this presentation are those of the author and do not reflect the official policy of the Department of Army/ Navy/Air Force, Department of Defense, or U.S. Government.