# Multimodal Phone-Call Analytics Identify Hospitalization Risk in Home Healthcare

ZHIHONG (SANDY) ZHANG, PHD, RN

DATA SCIENCE INSTITUTE, COLUMBIA UNIVERSITY, NEW YORK, UNITED STATES
SCHOOL OF NURSING, COLUMBIA UNIVERSITY, NEW YORK, UNITED STATES

# Acknowlegement

## Coauthors

- Pallavi Gupta, PhD,
- William Ho, MS,
- Yu-Wen Chen,
- Sasha Vergez, MS,
- Margaret McDonald, MS,
- Zoran Kostic, PhD,
- Julia Bell Hirschberg, PhD,
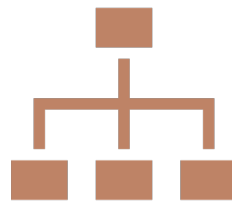- Maryam Zolnoori, PhD,
- Maxim Topaz, PhD

## Funding

# The Challenge: Early Risk Identification in Home Healthcare (HHC)

HHC patients have high rates of acute care utilization.

Traditional methods (self-reports, structured assessments) often miss early deterioration.

We need a **lightweight, continuous monitoring tool** that integrates into routine care.

# The Opportunity: Follow-ups phone calls

- **Verbal Content (Transcripts):** Captures explicit symptoms and health changes.

- **Non-Verbal Cues (Audio/Vocal Characteristics):** Offers rich, *non-explicit* information.

- **Hypothesis:** Verbal and non-verbal features from brief, follow-up calls can predict hospitalization/ED risk.

# Study objectives
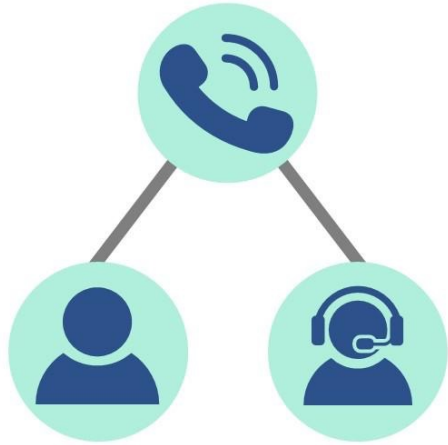
- To evaluate whether short, structured follow-up phone calls could be used to monitor hospitalization and ED visits among HHC patients.
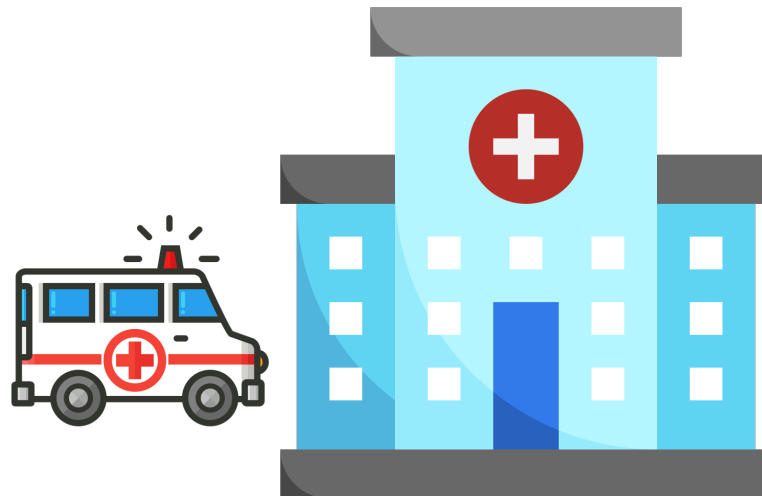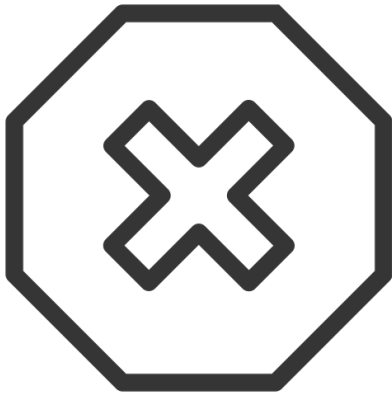
# Methods: Patients

- Patients were recruited from **VNS Health**, one of the largest nonprofit home healthcare agencies in the U.S.

- **Eligibility:** Patients who could communicate independently in English with HHC clinicians.

- **Exclusions:** Limited English proficiency or speech impairment (e.g., aphasia, Parkinson's disease).

# Phone call follow-ups

RAs conducted follow-up calls after home visits

Each call lasted <10 minutes

Up to 4 calls per patient
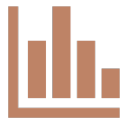
Average interval: 1 week between calls

Each call followed a structured 3-question protocol:

1) "How have you been feeling lately?"

2) "Any changes in your health over the past few days?"

3) "What symptoms have been bothering you lately?"

# Analytical Approach

Full transcript → Text-Based LLM: Qwen-2-7B-Instruct

Up to four calls per patient

- Leveraging Longitudinal LLM Architectures
- **Risk Stratification: LOW / MODERATE / HIGH**

30s-audio speech → Multimodal LLM: Qwen-2.5-Omni
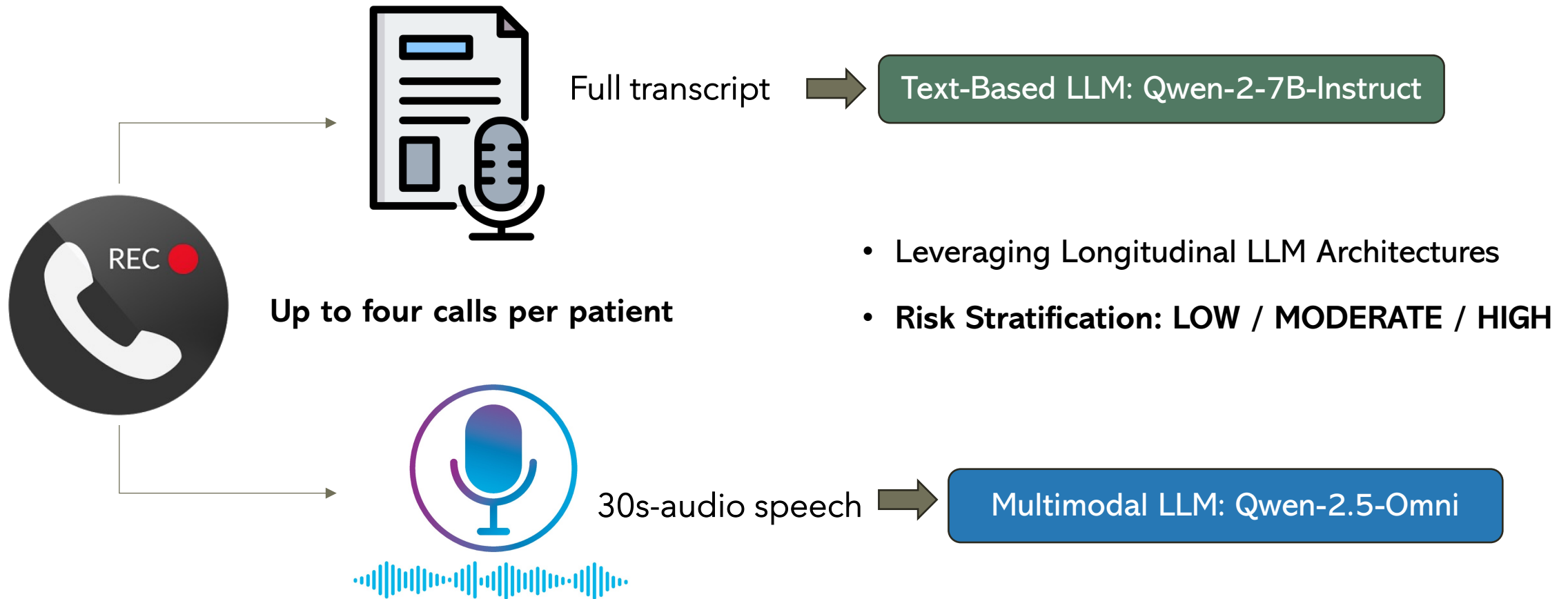
# Prompt examples

## AUDIO RISK ASSESSMENT (Example)

You are analyzing a **healthcare audio session** to predict **60-day hospitalization or ED visit risk.**

**Listen for key clinical indicators:**

- **Vocal stress:** strained, weak, or labored tone
- **Breathing:** shortness of breath, wheezing, or effortful speech
- **Energy level:** fatigue, low energy, or slowed pace
- **Pain cues:** groaning or sounds of discomfort
- **Cognitive status:** confusion, unclear, or disorganized speech
- **Emotional tone:** anxiety, sadness, or worry

**Assign a risk score (0.0–1.0):**

- **Low (0.00–0.33):** stable, healthy voice patterns
- **Moderate (0.34–0.66):** some concerning indicators
- **High (0.67–1.00):** multiple red flags or urgent concerns

## LONGITUDINAL RISK ASSESSMENT (Example)

You are a **clinical AI assistant** predicting **60-day hospitalization or ED visit risk** for home healthcare patients using **longitudinal call transcripts.**

Analyze all transcripts from **{num_sessions} sequential HHC calls** to assess overall risk trends.

**Key risk indicators:**

- **High (0.67–1.00):** acute symptoms (chest pain, SOB), recent falls, medication non-adherence, confusion, poor support, multiple red flags.
- **Moderate (0.34–0.66):** pain management issues, mild decline, some medication or caregiver concerns, intermittent symptoms.
- **Low (0.00–0.33):** stable condition, good adherence, strong support, improving symptoms, proactive self-management.

# Characteristic of patients

| Characteristic | | Total (N=180) | Event Group (N=44) | No Event Group (N=136) |
|---|---|---|---|---|
| **Age (mean ± SD)** | | 70.9 ± 14.0 | 72.0 ± 13.1 | 70.6 ± 14.2 |
| **Gender** | Female | 124 (68.9%) | 34 (77.3%) | 90 (66.2%) |
| | Male | 56 (31.1%) | 10 (22.7%) | 46 (33.8%) |
| **Ethnicity** | Non-Hispanic | 152 (84.4%) | 32 (72.7%) | 120 (88.2%) |
| | Hispanic | 16 (8.9%) | 5 (11.4%) | 11 (8.1%) |
| | Unknown | 12 (6.7%) | 7 (15.9%) | 5 (3.7%) |
| **Race** | White | 82 (45.6%) | 14 (31.8%) | 68 (50.0%) |
| | Black | 70 (38.9%) | 18 (40.9%) | 52 (38.2%) |
| | Other | 14 (7.8%) | 3 (6.8%) | 11 (8.1%) |
| | Unknown | 13 (7.2%) | 8 (18.2%) | 5 (3.7%) |
| | Asian | 1 (0.6%) | 1 (2.3%) | 0 (0.0%) |
| **Insurance (Medicare)** | Yes | 144 (80.0%) | 39 (88.6%) | 105 (77.2%) |
| | No | 36 (20.0%) | 5 (11.4%) | 31 (22.8%) |
| **Insurance (Medicaid)** | No | 123 (68.3%) | 24 (54.5%) | 99 (72.8%) |
| | Yes | 57 (31.7%) | 20 (45.5%) | 37 (27.2%) |
| **Insurance (dual)** | No | 137 (76.1%) | 28 (63.6%) | 109 (80.1%) |

# Characteristics of phone call recordings

| Recording Type | Count | Mean | Median | Std |
|---|---|---|---|---|
| Follow-up call 1 | 176 | 8:46 | 5:54 | 8:20 |
| Follow-up call 2 | 114 | 5:10 | 3:45 | 4:16 |
| Follow-up call 3 | 108 | 4:42 | 3:30 | 3:17 |
| Follow-up call 4 | 101 | 5:24 | 4:50 | 2:50 |

# Transcript-Predicted High-Risk Group Had Higher Hospitalization/ED Rates



Hospitalization Rates by Full Transcript
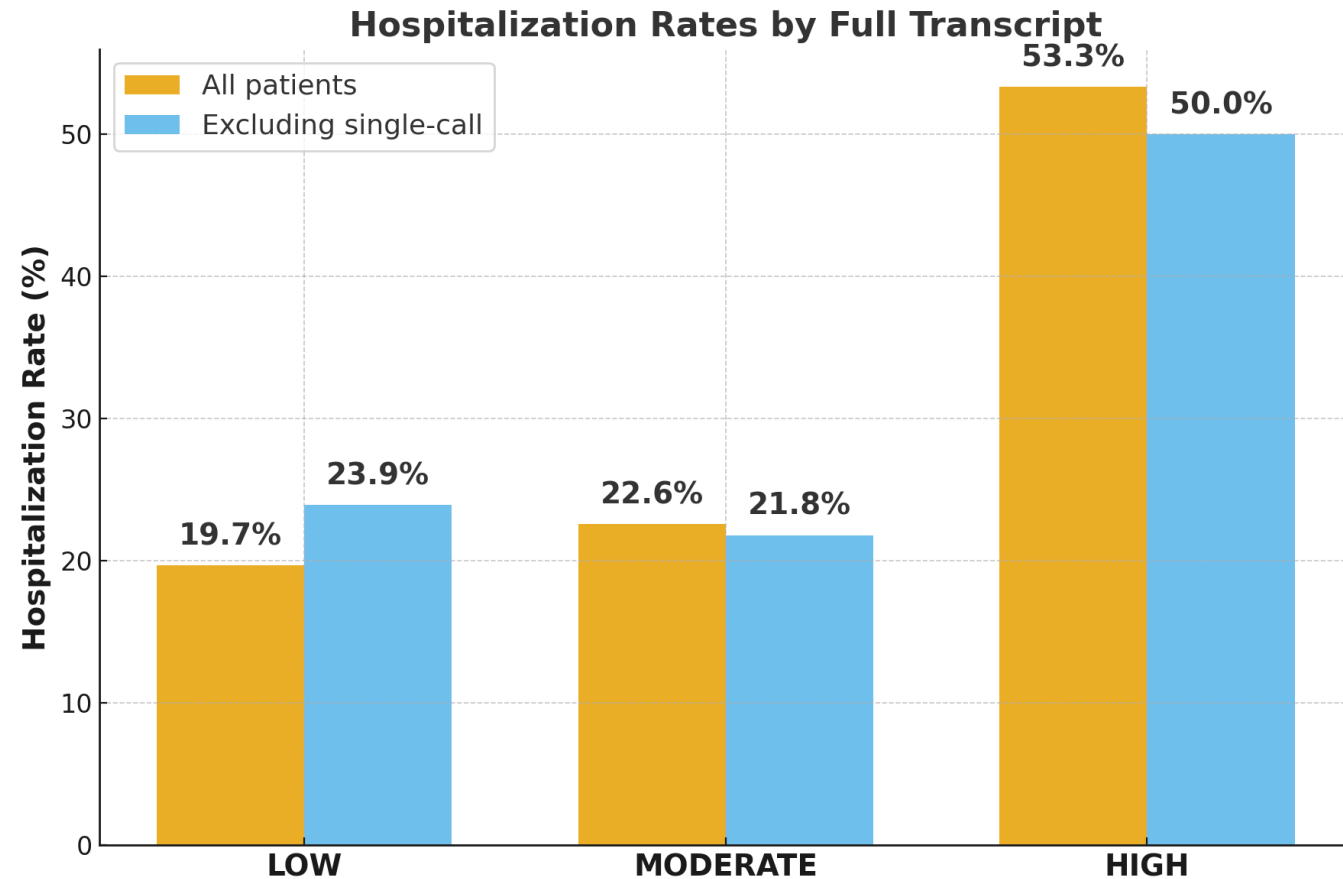
# Higher odds of hospitalization/ED visits in transcript-predicted high-risk group

## All patients

| Group | Category | N | Percent | OR | 95% CI | p-value |
|---|---|---|---|---|---|---|
| LOW (reference) | LOW | 76 | 19.7% | 1.0 | — | — |
| MODERATE vs LOW | MODERATE | 84 | 22.6% | 1.2 | 0.6 – 2.5 | 0.702 |
| HIGH vs LOW | HIGH | 15 | 53.3% | 4.6 | 1.5 – 14.8 | **0.019** |

## Patients excluding single calls

| Group | Category | N | Percent | OR | 95% CI | p-value |
|---|---|---|---|---|---|---|
| LOW (reference) | LOW | 46 | 23.9% | 1.0 | — | — |
| MODERATE vs LOW | MODERATE | 55 | 21.8% | 0.9 | 0.4 – 2.3 | 0.816 |
| HIGH vs LOW | HIGH | 14 | 50.0% | 3.2 | 0.9 – 11.1 | 0.095 |

# Audio-predicted high-risk group had higher hospitalization/ED rates



**Hospitalization Rates by 30-Second Speech**
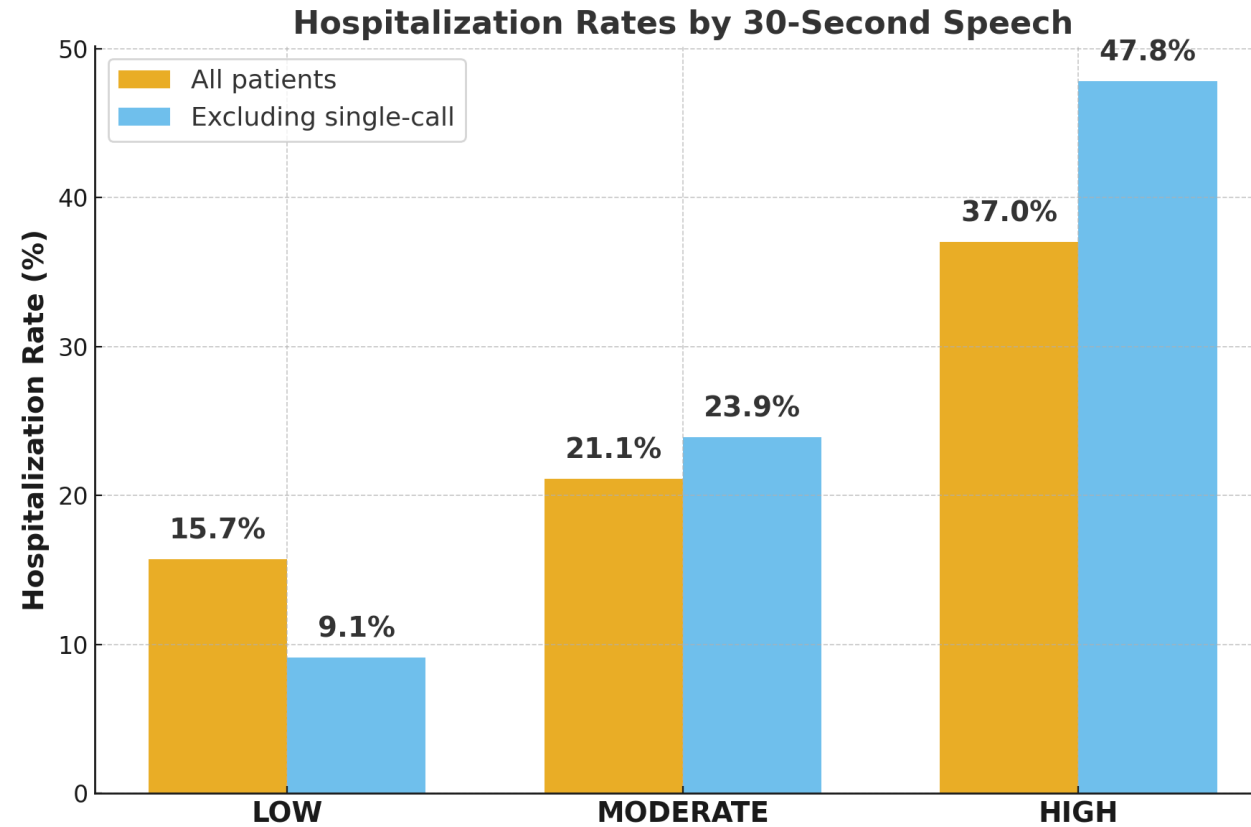
# Higher odds of hospitalization/ED visits in audio-predicted high-risk group

## All patients

| Group | Category | N | Percent | OR | 95% CI | p-value |
|---|---|---|---|---|---|---|
| LOW (reference) | LOW | 51 | 15.7% | 1.0 | — | — |
| MODERATE vs LOW | MODERATE | 76 | 21.1% | 1.4 | 0.6 – 3.7 | 0.496 |
| HIGH vs LOW | HIGH | 46 | 37.0% | 3.2 | 1.2 – 8.3 | **0.021** |

## Patients excluding single calls

| Group | Category | N | Percent | OR | 95% CI | p-value |
|---|---|---|---|---|---|---|
| LOW (reference) | LOW | 22 | 9.1% | 1.0 | — | — |
| MODERATE vs LOW | MODERATE | 67 | 23.9% | 3.1 | 0.7 – 14.9 | 0.220 |
| HIGH vs LOW | HIGH | 23 | 47.8% | 9.2 | 1.7 – 48.6 | **0.007** |

# Verbal and non-verbal cues from follow-up phone calls signal early clinical deterioration

Full-transcript model captured meaningful **verbal patterns** from call content.

30-second speech audio capture **subtle vocal tone, rhythm, and speech dynamics** over time.

Both **verbal and non-verbal cues** from brief, structured calls can signal **early clinical deterioration** in HHC patients.

# Conclusions

Brief, structured follow-up calls capture valuable data for hospitalization risk stratification.

Findings support integrating lightweight, multimodal call analytics into routine HHC monitoring to enable proactive and timely interventions.

Zhihong (Sandy) Zhang, PhD, RN
zz3238@columbia.edu

VNS Health