



# Diagnostic and Management Applications of ChatGPT in Otolaryngology Clinical Scenarios

Roy W. Qu MD, Uneeb Qureshi MD, Garrett Petersen BS, Steve C. Lee MD, PHD

## Introduction & Background

- Artificial Intelligence (AI) applications are increasingly pervading the healthcare space.
- Understanding the utility of AI is important for both clinicians, and their patients, given these models are increasingly publically available.
- The efficacy of AI and Language Learning Models (LLM) in the field of otolaryngology in developing assessment and plans with that of human physicians was assessed with various clinical vignettes.

## Hypothesis

- We hypothesized that ChatGPT would not surpass physicians in terms of diagnostic and treatment plan development

## Methods

- ChatGPT 4.0 was queried for diagnoses and management plans for 20 physician-written clinical vignettes, containing medical jargon, in the field of otolaryngology. Attending physicians rated difficulty and responses to the clinical vignettes on a 5-point Likert scale.
- Summary statistics were calculated. Univariate ordinal regression was then performed between vignette difficulty and quality of the diagnoses and management plans.

## Results

- ChatGPT provides accurate differential diagnoses and reasonable treatment plans.
- It remains unclear how well it would perform if provided with real-world, open-ended free-form histories, as well as complete physical exams which would often contain irrelevant, extraneous, or contradictory information.
- Limitations of ChatGPT and other LLMs stems from their proprietary databases. At the time of this publication it lacked data after September 2021, did not learn from prior experience, and often required prompting for clarification.
- Other risks include artificial hallucinations (i.e. “produce content that is nonsensical or untruthful in relation to certain sources”), e.g. fabricated references when asked for sources

6. A 40 year-old female comes in with a 12 month history of progressive right hearing loss. It is associated with a high frequency non-pulsatile tinnitus and intermittent vertigo. She denies otalgia and otorrhea. Recently, she has noticed the right-side of her face drooping over the past 3 months. On physical exam, her external auditory canals and tympanic membranes are clear. There is no middle ear effusion. Her tuning fork exam lateralizes to the left, and air conduction is greater than bone conduction bilaterally. She has a House-Brackmann grade 3 right facial paralysis.

- 6
1. Vestibular schwannoma (acoustic neuroma)
  2. Ramsay Hunt Syndrome (herpes zoster oticus)
  3. Ménière's disease
  4. Temporal bone fracture
  5. Autoimmune inner ear disease

6 Based on the patient's history and presentation, the most likely diagnosis is a vestibular schwannoma (acoustic neuroma). Here's a suggested treatment plan:

Consults – neuro-otologist, NSGY, audiologist  
Radiographic studies – MRI head w/ contrast  
Treatment options – observation (small tumors with mild symptoms), stereotactic radiosurgery (serviceable hearing and tumor < 3cm), surgical resection (large tumors with significant symptoms or when hearing preservation is not possible)  
Follow up – facial nerve rehabilitation, neuro-otology, NSGY

## Conclusions

- ChatGPT and other LLMs are promising technologies for patient and provider use
- Assessments and plans for easy to moderately difficult clinical scenarios seem to be acceptable to current standards of care.
- Users must be aware of the challenges and limitations of any AI or LLM technology before implementing it in clinical practice. As evidenced by recent controversies.
- Physicians, medical societies, and patients, among other important stakeholders, should be involved in the development and application of these nascent technologies.

Likert score	Dependent: Diagnosis score			Dependent: Treatment score			Dependent: Treatment score		
	OR	95% CI	p Value	OR	95% CI	p Value	OR	95% CI	p Value
1	2.01	(0.275, 14.30)	.495	0.482	(0.046, 5.00)	.542	0.381	(0.060, 2.44)	.307
2	0.853	(0.121, 5.99)	.874	0.273	(0.027, 2.80)	.274	0.330	(0.110, 1.00)	.050
3	0.865	(0.106, 7.03)	.892	0.868	(0.071, 10.70)	.912	0.289	(0.082, 1.01)	.052
4	0.660	(0.087, 5.00)	.162	0.232	(0.021, 2.53)	.231	0.301	(0.170, 0.533)	<.001
5	1	-	-	1	-	-	1	-	-