Voice Acoustic Measures of Depression Severity and Treatment Response Collected with IVR

Heidi K. Moore, PhD
Healthcare Technology Systems, Inc.
August 7, 2006
A Brief History
Clinical Computing

First computer-automated patient interview
U Wisconsin – Madison
1967

LINC 8
Magnetic tape IO
8 inch CRT
4x5 matrix chars
10 cps teletype
"Diagnostic computers don't make mistakes, Mr. Pomeroy. You have Dutch elm disease."
The Birth of ePRO

electronic

Patient

Reported

Outcomes
Standardized clinical assessments that electronically obtain information from patients

<table>
<thead>
<tr>
<th>Depression</th>
<th>Insomnia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxiety</td>
<td>Pain</td>
</tr>
<tr>
<td>Substance use/abuse</td>
<td>OCD</td>
</tr>
<tr>
<td>Diabetes</td>
<td>Asthma</td>
</tr>
<tr>
<td>Hypertension</td>
<td>Quality of life</td>
</tr>
</tbody>
</table>
“You have reached the Complaint Department. If you are moderately angry, Press One. If you are extremely angry, Press Two…”
An Examination of 26,168 Hamilton Depression Rating Scale Scores Administered via Interactive Voice Response Across 17 Randomized Clinical Trials

Heidi K. Moore, PhD,* James C. Mundt, PhD,* Jack G. Modell, MD,† Heidi E. Rodrigues, BS,‡¶ David J. DeBrota, MD.§ James J. Jefferson, MD,* and John H. Greist, MD*

**FDA statement regarding computerized instruments**

1. We would accept self-report in general as an approach to assessing symptom severity in drug trials focusing on major depressive disorder in outpatients. This is true, whether that approach involves IVR technology, paper and pencil, or other means.

2. HAM-D done by IVR would be one acceptable approach to obtaining these assessments.

3. Regarding the Inventory of Depressive Symptoms (IDS) and the Quick Inventory of Depressive Symptoms (QIDS) in their various forms, that is, IDS-C, IDS-SR, QIDS-C, QIDS-SR, QIDS-IVR, we would also find any of these acceptable means to assessing symptom severity in drug trials focusing on major depressive disorder in outpatients.

Regarding how sponsors should proceed if the intend to use one of these assessments, they should submit protocols that include their preferred assessment. If they need prior confirmation that a particular instrument would be acceptable, they could e-mail me. However, if it is one of the above, my response would be the same.
Quick Inventory of Depressive Symptomatology – Adolescent version (QIDS-A IVR)

- Adolescent adaptation of popular IVR adult depression assessment
- Modified wording to be easily understood by adolescents (“focus attention” to “pay attention”)
- Speech-enabled assessment allowed more efficient assessment
- Leveled response burden to discourage patterned responding
- Recognized “Yes” “No” and whole numbers from 0 -24
- Compared to paper-based and clinician QIDS and field standard depression assessment (CDRS-R) in pilot study (n= 27)
Relationship of QIDS-A IVR to Paper QIDS

$r = .82$
Relationship of QIDS-A IVR to Clinician QIDS

$r = .95$
Relationship of QIDS-A IVR to CDRS-R

$r = .76$
QIDS-A IVR Pilot Results

• Average completion time 6:31 (SD= 0:41)

• Assessment correlated significantly and positively with other measures of depression

• Adolescents responded positively to speech-enabled clinical assessment; only two instances of defaulting to DTMF
Objective Biomarkers of Depression in Vocal Acoustics

- Speech from depressed patients is perceptibly different in pitch, loudness, speaking rate, and articulation.
- Clinicians attend to "How" patients speak as well as "What" they say.
- Production of speech reflects underlying neurophysiologic functioning.
- Vocalization may be under less conscious control than selection and execution of speech production.
Remote capture of human voice acoustical data by telephone: A methods study

MICHAEL S. CANNIZZARO¹, NICOLE REILLY¹, JAMES C. MUNDT², & PETER J. SNYDER¹,³

¹Voice Acoustics Laboratory, WW Clinical Technology (CNS), Pfizer Global Research & Development, Groton Laboratories, Groton, CT, USA, ²Healthcare Technology Systems, Inc., Madison, WI, USA, and ³Department of Psychology, University of Connecticut, Storrs, CT, USA

(Received 3 February 2004; accepted 7 June 2004)
Speech Elicitation Tasks

- Extemporaneous “Free” speech
  - Describe emotional, physical, functional experiences during the past week

- Automatic “Rote” speech
  - Count from 1 to 20
  - Recite the alphabet
  - Read ‘The Grandfather Passage’

- Held vowels
  - /a/, /i/, /u/, and /ae/ held for 5 seconds

- Diodochokineses
  - /pa ta ka/ repeated for 5 seconds
Which Aspects of Speech are Measured?

- Pitch variability $F_0$, $F_1$, and $F_2$ (COV)
- Total recording duration
- Vocalization time
- Pause characteristics (number, length, SD)
- Percent pause time
- Vocalization/pause ratio
- Speaking rate (syllables per second)
Digital signal processing of running speech generates a quantifiable measurement of motor ability.
Tryout Study Design

- 35 patients starting treatment for depression
- Referred by treating physicians
- 20 women, 15 men (20 to 68 years old)
- Six weeks of follow-up
- Depression assessed weekly via IVR
- Speech samples collected weekly
Patients’ Response to Treatment

Depression severity reduced by 50% or more

Week 1
- 90.9% (Blue)
- 9.1% (Green)

Week 2
- 85.3% (Blue)
- 14.7% (Green)

Week 3
- 70.0% (Blue)
- 30.0% (Green)

Week 4
- 81.8% (Blue)
- 18.2% (Green)

Week 5
- 55.2% (Blue)
- 44.8% (Green)

Week 6
- 59.4% (Blue)
- 40.6% (Green)
Analysis of Vocal Acoustics

More severe depression =

↓ Pitch Variability
↑ Recording Length
↑ Total pause time
↑ Pause variability
↑ Percent time pausing
↓ Vocalization/pause ratio
↓ Speaking rate (syllables/sec)
After Six Weeks of Treatment

Vocal Acoustic Change Associated with Treatment Response

- Pitch Variability: $p = 0.02$
- Speaking Rate: $p = 0.002$
- Number Pauses: $p = 0.03$
- Pause Time: $p = 0.05$
- Recording Length: $p = 0.02$
Conclusions

• Automation of speech collection procedures over the telephone to obtain reliable and valid voice acoustic metrics is clearly feasible.

• Vocal acoustic measures of pitch variability and speech production are significantly influenced by depression and response to treatment.

• Reduced speaking time for responders was due to fewer pauses, less pause time, and faster speaking rates, not changes in vocalization time.

• Results of this study are consistent with previous research.

• Currently studying ability of IVR vocal acoustics to detect early response to antidepressant treatment in a randomized clinical trial.
Voice acoustic measures of depression severity and treatment response collected via interactive voice response (IVR) technology

James C. Mundt\textsuperscript{a,\ast}, Peter J. Snyder\textsuperscript{b,c}, Michael S. Cannizzaro\textsuperscript{b,d}, Kara Chappie\textsuperscript{b}, Dayna S. Geralts\textsuperscript{a}

\textsuperscript{a}Healthcare Technology Systems, Inc., Healthcare Technology Systems, Inc., 7617 Mineral Point Road, Suite 300, Madison, WI 53717, USA
\textsuperscript{b}Pfizer Global Research & Development, Groton Laboratories, Groton, CT, USA
\textsuperscript{c}Department of Psychology, University of Connecticut, Storrs, CT, USA
\textsuperscript{d}Department of Communication Sciences, University of Vermont, Burlington, VT, USA

Received 10 January 2006; received in revised form 27 March 2006; accepted 4 April 2006
QIDS-A IVR research support by NIMH (RO1MH068851, AJ Rush, Principle Investigator) and by GlaxoSmithKline, Inc.

Vocal acoustics research support by NIMH (R43MH68950 JC Mundt, Principle Investigator) and by Pfizer, Inc.