Correlation of natural language assessment results with health-related quality of life in adult glioma patients

Simon G. Ammanuel, BS,1 Nyle C. Almeida, BS,1 Garret Kurteff, BA,1,2 Sofia Kakaizada,1 Annette M. Molinaro, PhD,1 Mitchel S. Berger, MD,1 Edward F. Chang, MD,1,3 and Shawn L. Hervey-Jumper, MD1

1Department of Neurological Surgery and 1Center for Integrative Neuroscience, University of California, San Francisco, California; and 3Department of Communication Sciences & Disorders, University of Texas, Austin, Texas

OBJECTIVE Impairments of speech are common in patients with glioma and negatively impact health-related quality of life (HRQoL). The benchmark for clinical assessments is task-based measures, which are not always feasible to administer and may miss essential components of HRQoL. In this study, the authors tested the hypothesis that variations in natural language (NL) correlate with HRQoL in a pattern distinct from task-based measures of language performance.

METHODS NL use was assessed using audio samples collected unobtrusively from 18 patients with newly diagnosed low- and high-grade glioma. NL measures were calculated using manual segmentation and correlated with Quality of Life in Neurological Disorders (Neuro-QoL) outcomes. Spearman’s rank-order correlation was used to determine relationships between Neuro-QoL scores and NL measures.

RESULTS The distribution of NL measures across the entire patient cohort included a mean ± SD total time speaking of 11.5 ± 2.20 seconds, total number of words of 27.2 ± 4.44, number of function words of 10.9 ± 1.68, number of content words of 16.3 ± 2.91, and speech rate of 2.61 ± 0.20 words/second. Speech rate was negatively correlated with functional domains (rho = −0.62 and p = 0.007 for satisfaction with social roles; rho = −0.74 and p < 0.001 for participation in social roles) but positively correlated with impairment domains (rho = 0.58 and p = 0.009 for fatigue) of Neuro-QoL.

CONCLUSIONS Assessment of NL at the time of diagnosis may be a useful measure in the context of treatment planning and monitoring outcomes for adult patients with glioma.

https://thejns.org/doi/abs/10.3171/2021.1.JNS203387

KEYWORDS glioma; natural language; aphasia; task-based language testing; oncology
factors, such as the prevalence of other nonaphasic neurological impairments and chronic fatigue.\textsuperscript{16} Williamson et al. conducted a study of stroke patients and showed that WAB scores did not correlate with patients’ quality-of-life outcomes.\textsuperscript{17} Therefore, it may be optimal for clinicians to have the ability to establish risk and predict a patient outcome on the basis of parameters outside task performance. Patterns of NL may be of prognostic significance in clinical populations; however, they have not been studied in adult glioma patients.

NL is the instinctual written or spoken language profile that an individual has developed throughout his or her lifetime.\textsuperscript{18–22} Recent psycholinguistic analyses suggest that change in patterns of NL may be an accurate measure for assessing clinical and behavioral conditions. Changes in NL may include shifts in specific function words or alterations in total language output. NL analysis of speech has been shown to determine changes in cognitive status and possibly detect early stages of dementia.\textsuperscript{23} In Parkinson’s disease patients, a slower speech rate was correlated with worsening disease progression.\textsuperscript{24} Variations in NL structure may offer unique behavioral indicators of short-term health-related outcomes in glioma patients. Analysis of the aforementioned information-rich components of NL processing may be accomplished either manually or via machine learning algorithms. In this study, we tested the hypothesis that baseline NL measures of adult patients with glioma would correlate with preoperative scores on functional and impairment HRQoL domains in patterns distinct from task-based measures of language performance.

**Methods**

The study received institutional review board approval from the University of California, San Francisco Committee for Human Research. All patients provided written informed consent to participate after the study’s nature and design were fully described and related questions had been answered.

Consecutive adult patients aged 18–85 years at diagnosis with glioma completed initial language assessments. All tumors were within the frontal, parietal, or temporal lobes and diagnosed between August 2017 and September 2018. Only patients with dominant-hemisphere gliomas were included in this study, with the hemisphere of language dominance determined on the basis of cortical activation during picture-naming and text-reading language tasks that were administered with magnetoencephalography.\textsuperscript{25} The control group included patients with benign intracranial lesion (e.g., meningioma, colloid cyst). Patients were excluded from this study if they 1) had a prior cranial operation or traumatic brain injury, 2) were not primary English speakers, 3) had clinical or radiographic evidence of ongoing seizure activity, and/or 4) had baseline impaired cognition that rendered them unable to complete QAB language assessments. Language tasks included 1) picture naming, 2) text reading, 3) auditory naming, 4) syntax, and 5) 4-syllable word repetition.

For each patient, NL measures were analyzed from an average of 10 separate 30- to 60-second audio samples (10 minutes of audio samples per participant) obtained with a 60-minute cognitive assessment battery conducted by an independent study coordinator on a single day. The audio recordings were manually transcribed by one of the authors (S.G.A.), and NL coding was done using these transcriptions. NL output measures included total time speaking, total number of words, number of function words, number of content words, and speech rate, which is the ratio of total number of words to total time speaking.\textsuperscript{26}

Function words were defined as words that provided structural or grammatical relationships with other words. Content words were interpreted as words that name objects of reality and their qualities. Audio files were coded manually by coauthors blinded to clinical outcome. QAB was administered and scored by a speech pathologist or trained research assistant according to an established protocol.\textsuperscript{16} A score of 4 on QAB was a correct answer; a score of 3 was a correct answer that was delayed > 3 seconds or self-corrected; a score of 2 was an answer where at least half the phonemes were correct; a score of 1 was an answer that was incorrect but somewhat related to the target; and a score of 0 was an unrelated response within 6 seconds or no answer.

HRQoL was assessed using adult version 1.1 of the Quality of Life in Neurological Disorders (Neuro-QoL) assessment tool.\textsuperscript{8} Neuro-QoL is a set of questionnaires that addresses common issues in neurological diseases. Each domain is stratified into two categories: function (cognition, participation and satisfaction with social roles and activities, upper- and lower-extremity function, and positive affect) and impairment (stigma, sleep disturbance, fatigue, emotional dyscontrol, depression, and anxiety).\textsuperscript{8} Specifically, for this tool, anxiety was defined as unpleasant thoughts or feelings related to fear, helplessness, worry, and hyperarousal. Depression was defined as self-reported feelings of loss, hopelessness, negative mood, decreased positive affect, or negative views of self. Fatigue was described as self-reported sensations ranging from tiredness to an overwhelming, debilitating, and sustained sense of exhaustion that decreases one’s capacity for physical, functional, social, and mental activities. Upper-extremity function involved one’s ability to carry out various activities involving digital, manual, and reach-related functions (including fine motor and self-care functions). Lower-extremity function included one’s ability to carry out various activities involving the trunk region, body movement, ambulation, balance, or endurance. The cognitive functional domain was used to assess perceived difficulties in cognitive abilities or the application of such abilities to everyday tasks. Emotional dyscontrol consisted of a set of disease and treatment-related manifestations, including disinhibition, emotional lability, irritability, impatience, and impulsiveness. Positive affect and well-being included aspects of a person’s life related to a sense of well-being, satisfaction, or overall sense of purpose and meaning. Sleep disturbance included perceptions of sleep quality, sleep depth, and restoration associated with sleep. The ability to participate in social roles and activities was the degree of involvement in one’s usual social roles, activities, and responsibilities, including those related to work, family, and friends. At the same time, satisfaction with social roles and responsibilities was defined as satisfaction with involvement in one’s usual social roles,
activities, and responsibilities. Stigma involved perceptions of self and publicly enacted negativity, prejudice, and discrimination due to disease.\textsuperscript{8} Neuro-QoL was administered and scored by a trained independent research assistant according to an established protocol.\textsuperscript{8} Each assessment was scored with normalized mean t-scores for each domain and standardized to a score of 50. Scores higher than 60 on the functional domains and lower than 50 on the impairment domains were considered desirable. Impairment was assessed in the study population, and patients with scores > 1 SD beyond the normative mean were deemed to have impaired function.

Descriptive statistics were used to summarize group characteristics, NL performance, and Neuro-QoL scores. Comparison of NL parameters between the control and glioma patients was performed using the nonpaired t-test. Spearman correlation coefficients were assessed for each NL and QAB measures assessed, we applied Bonferroni correction to control for multiple comparisons. Thus, to adjust for the 5 NL measures and 5 QAB tasks, Spearman correlation coefficients were considered statistically significant at p < 0.01. Data were reported as mean and standard error. The alpha level for significance was set at 0.05. Statistical analysis was performed using MATLAB version R2019a (The MathWorks, Inc.).

**Results**

This study included 23 adult patients: 18 with newly diagnosed grade II, III, and IV glioma based on the 2016 WHO diagnostic criteria, and 5 in the control group (4 with meningioma and 1 with colloid cyst). The data set included 230 audio samples, totaling 13,040 seconds of unstructured speech, that were manually coded for 5 NL measures. Demographic and tumor characteristics are listed in Table 1.

We first set out to understand and compare the distributions of NL measures across the glioma and control cohorts. This included a mean ± SD total time speaking of 11.5 ± 2.20 seconds for the glioma cohort versus 8.94 ± 1.3 seconds for the control cohort (p = 0.56), total number of words of 27.2 ± 4.44 words versus 28.92 ± 1.95 words (p = 0.84), number of function words of 10.9 ± 1.68 words versus 13.28 ± 1.01 words (p = 0.48), number of content words of 16.3 ± 2.91 words versus 15.64 ± 2.67 words (p = 0.91), and speech rate of 2.61 ± 0.20 words/second versus 3.47 ± 0.30 words/second (p = 0.049). The distributions of the NL measures between the two groups are shown in Fig. 1.

HRQoL was then measured using Neuro-QoL. Functional domain assessments included a mean t-score for cognition of 41.7 ± 1.43, satisfaction with social roles of 44.3 ± 1.06, upper-extremity function of 41.7 ± 2.64, positive affect of 53.3 ± 1.68, lower-extremity function of 48.4 ± 2.41, and participation in social roles of 43.6 ± 1.23. Across the Neuro-QoL impairment domain, patients had a mean t-score for stigma of 50.3 ± 1.80, sleep disturbance of 53.3 ± 1.74, fatigue of 47.4 ± 2.29, emotional dyscontrol of 49.8 ± 2.75, depression of 49.9 ± 1.93, and anxiety of 55.98 ± 2.09. Figure 2 shows the distributions of Neuro-QoL scores across patients.

Decreased HRQoL was defined as > 1 SD from established normative values according to an established protocol.\textsuperscript{8} Prevalence of decreased HRQoL was highest for cognition (50% of patients), upper-extremity function (50%), anxiety (33.3%), lower-extremity function (27.8%), participation in social roles (27.8%), satisfaction with social roles (22.2%), sleep disturbance (22.2%), and emotional dyscontrol (22.2%; Table 2).

|TABLE 1. Demographic and clinical characteristics of the glioma cohort|

<table>
<thead>
<tr>
<th>Patient No.</th>
<th>Age (yrs)</th>
<th>Sex</th>
<th>Dominant Hand</th>
<th>Tumor Side</th>
<th>Tumor Location</th>
<th>WHO Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>56</td>
<td>M</td>
<td>Lt</td>
<td>Rt</td>
<td>Frontal</td>
<td>IV</td>
</tr>
<tr>
<td>2</td>
<td>60</td>
<td>M</td>
<td>Lt</td>
<td>Lt</td>
<td>Temporal</td>
<td>IV</td>
</tr>
<tr>
<td>3</td>
<td>69</td>
<td>F</td>
<td>Rt</td>
<td>Lt</td>
<td>Temporal</td>
<td>IV</td>
</tr>
<tr>
<td>4</td>
<td>41</td>
<td>F</td>
<td>Rt</td>
<td>Lt</td>
<td>Temporal</td>
<td>II</td>
</tr>
<tr>
<td>5</td>
<td>60</td>
<td>M</td>
<td>Rt</td>
<td>Lt</td>
<td>Frontal</td>
<td>IV</td>
</tr>
<tr>
<td>6</td>
<td>62</td>
<td>F</td>
<td>Rt</td>
<td>Lt</td>
<td>Temporal</td>
<td>IV</td>
</tr>
<tr>
<td>7</td>
<td>65</td>
<td>M</td>
<td>Rt</td>
<td>Lt</td>
<td>Temporal</td>
<td>IV</td>
</tr>
<tr>
<td>8</td>
<td>66</td>
<td>F</td>
<td>Rt</td>
<td>Lt</td>
<td>Frontal</td>
<td>IV</td>
</tr>
<tr>
<td>9</td>
<td>26</td>
<td>F</td>
<td>Rt</td>
<td>Lt</td>
<td>Parietal</td>
<td>II</td>
</tr>
<tr>
<td>10</td>
<td>41</td>
<td>F</td>
<td>Rt</td>
<td>Lt</td>
<td>Frontal</td>
<td>IV</td>
</tr>
<tr>
<td>11</td>
<td>64</td>
<td>F</td>
<td>Rt</td>
<td>Lt</td>
<td>Temporal</td>
<td>IV</td>
</tr>
<tr>
<td>12</td>
<td>57</td>
<td>M</td>
<td>Rt</td>
<td>Lt</td>
<td>Temporal</td>
<td>IV</td>
</tr>
<tr>
<td>13</td>
<td>57</td>
<td>M</td>
<td>Rt</td>
<td>Lt</td>
<td>Temporal</td>
<td>IV</td>
</tr>
<tr>
<td>14</td>
<td>60</td>
<td>F</td>
<td>Rt</td>
<td>Lt</td>
<td>Temporal</td>
<td>IV</td>
</tr>
<tr>
<td>15</td>
<td>26</td>
<td>M</td>
<td>Rt</td>
<td>Lt</td>
<td>Frontal</td>
<td>II</td>
</tr>
<tr>
<td>16</td>
<td>43</td>
<td>F</td>
<td>Lt</td>
<td>Rt</td>
<td>Frontal</td>
<td>III</td>
</tr>
<tr>
<td>17</td>
<td>49</td>
<td>M</td>
<td>Rt</td>
<td>Lt</td>
<td>Frontal</td>
<td>IV</td>
</tr>
<tr>
<td>18</td>
<td>64</td>
<td>F</td>
<td>Lt</td>
<td>Rt</td>
<td>Temporal</td>
<td>IV</td>
</tr>
</tbody>
</table>
We then set out to determine whether NL measures were correlated with results on the functional or impairment domains of Neuro-QoL and whether these correlations differed from those of task-based language assessments. The Bonferroni correction was applied because NL measures included comparisons of multiple variables within the same family. Several Neuro-QoL domains were significantly correlated with speech rate, including satis-

![Image of graphs](image-url)
TABLE 2. Decreased HRQoL in the glioma cohort (n = 18)

<table>
<thead>
<tr>
<th>Domain</th>
<th>No. (%) w/ Decreased HRQoL*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Functional</td>
<td></td>
</tr>
<tr>
<td>Cognition</td>
<td>9 (60.0)</td>
</tr>
<tr>
<td>Satisfaction w/ social roles</td>
<td>4 (22.2)</td>
</tr>
<tr>
<td>Upper-extremity function</td>
<td>9 (60.0)</td>
</tr>
<tr>
<td>Positive affect</td>
<td>1 (5.6)</td>
</tr>
<tr>
<td>Lower-extremity function</td>
<td>5 (27.8)</td>
</tr>
<tr>
<td>Participation in social roles</td>
<td>5 (27.8)</td>
</tr>
<tr>
<td>Impairment</td>
<td></td>
</tr>
<tr>
<td>Stigma</td>
<td>2 (11.1)</td>
</tr>
<tr>
<td>Sleep disturbance</td>
<td>4 (22.2)</td>
</tr>
<tr>
<td>Fatigue</td>
<td>2 (11.1)</td>
</tr>
<tr>
<td>Emotional dyscontrol</td>
<td>4 (22.2)</td>
</tr>
<tr>
<td>Depression</td>
<td>1 (5.6)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>6 (33.3)</td>
</tr>
</tbody>
</table>

* Defined as a score of 50 or less for functional domains, or a score of 50 or greater for impairment domains.

This study identified a relationship between NL speech rate and distinct HRQoL measures. Functional and cognitive measures, such as assessments focused on language processing, are increasingly recognized for their importance.27 Task-based measures represent the vast majority of functional assessments. However, they are not always feasible to administer and may not fully describe all components of speech.

NL is the written and spoken language profile that an individual develops throughout life.18–22 Generally speaking, NL has seven components: pragmatics, phonology, phonetics, morphology, lexicon, syntax, and semantics.30,28 Function words, including pronouns and adverbs, are generated automatically as opposed to consciously created meaning words.29 Function words by themselves have no semantic purpose, but together they provide a syntactic structure. Patterns of NL shift in response to disease; therefore, they may be useful measures of patient outcomes.8,29 This is the first study of NL in the adult neurooncology setting.

Patient-reported HRQoL outcomes are essential tools used to understand disease in clinical practice. Patients with low- and high-grade glioma experience a wide range of neurological symptoms because the disease has both oncological and neurological ramifications, both of which impact quality of life. Therefore, when considering functional measures such as NL processing, their relationships with HRQoL measures are central. In this study, our glioma cohort had a significantly lower speech rate than our control cohort. In our glioma cohort, there was a negative relationship between speech rate and both patient satisfaction and participation in social roles (functional domains), as well as a positive relationship between speech rate and fatigue (impairment domain). Satisfaction and participation in social roles are essential components of HRQoL that are otherwise difficult to assess with measures besides patient-reported outcomes. The fact that faster speech rate was associated with decreased t-scores, and therefore considerably more distress in social settings, may reflect the underlying state of the fluent versus dysnomic subtypes of aphasia that are seen in patients when glioma is diagnosed. Furthermore, it remains unknown whether the correlation between faster speech rate and decreased t-scores for fatigue (i.e., worse fatigue) was due to a cause or effect.

Taken together, these findings suggest that baseline NL measures are correlated with select Neuro-QoL scores obtained at the time of diagnosis. In particular, the correlation strength of speech rate with quality-of-life outcomes in this study is consistent with the findings of previous research on various other neurological diseases. Several studies have demonstrated that worsening speech disturbances is associated with disease progression and deteriorating quality of life in patients with Parkinson’s disease.24,30,31 Friedova et al. also showed that decreased articulation rate was associated with impairment in information-processing speed in patients with multiple sclerosis.32 Cordella et al. has demonstrated that speech rate is an excellent measure for diagnosing and differentiating subtypes of primary progressive aphasia.33

By its nature, glioma integrates within functional language networks, resulting in varying degrees of network disruption that can cause a variety of symptoms and lead to impairment in HRQoL status. This study’s findings highlight the potential use of NL, in particular speech rate, as a useful measure given its correlation with HRQoL. Although this preliminary study was focused on patient assessments performed during early-stage disease, these results may be particularly meaningful for longitudinal symptom monitoring. Several studies have shown that fatigue and social functioning have prognostic value for a patient’s overall survival.34–36 Oncological interventions, such as resection, chemoradiation, corticosteroids, and antiepileptic drugs, are known to impact HRQoL to varying degrees.37,38 NL measures such as speech rate may be useful markers of disease or symptom trajectory.

The limitations of the current study include the retrospective review of prospectively collected data, single-institution study design, and small sample size. Furthermore, we chose to focus on HRQoL and NL measures at the time of diagnosis; therefore, further study is warranted to assess for correlations throughout the course of disease. The inception of machine learning algorithms for the systematic extraction of NL measures may also present some advantages over a manual NL system, which was used in this study.4,29

Conclusions

This research identified a relationship between NL speech rate and distinct HRQoL measures. This is the first study to assess NL measures as a functional endpoint in the adult neurooncology setting.
Acknowledgments

Dr. Hervey-Jumper receives funding from the National Institute of Neurological Disorders and Stroke (K08NS110919-01), National Cancer Institute (SP50CA97257-15), Robert Wood Johnson Foundation (74259), and LoGlio Collaborative.

References


Disclosures
The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper.

Author Contributions
Conception and design: Ammanuel, Hervey-Jumper. Acquisition of data: Ammanuel, Almeida, Kakaiizada, Hervey-Jumper. Analysis and interpretation of data: Ammanuel, Hervey-Jumper. Drafting the article: all authors. Critically revising the article: all authors. Reviewed submitted version of manuscript: all authors. Analysis and interpretation of data: Ammanuel, Almeida, Kakaizada, Hervey-Jumper. Acquiring, administrative/technical/material support: Ammanuel. Study supervision: Ammanuel, Hervey-Jumper.

Supplemental Information
Previous Presentations
The abstract was previously presented at the 2019 Annual Meeting of the Congress of Neurological Surgeons, San Francisco, California, October 19–23, 2019.

Correspondence
Shawn L. Hervey-Jumper: University of California, San Francisco, CA. shawn.hervey-jumper@ucsf.edu.