Critical Review: Is diffusion tensor tractography of the arcuate fasciculus a valid anatomical method of predicting prognosis in patients with aphasia?

Kerry Hollingsworth
M.Cl.Sc (SLP) Candidate
Western University: School of Communication Sciences and Disorders

The published body of literature investigating the clinical utility of diffusion tensor tractography imaging of the arcuate fasciculus to predict language outcomes in patients with aphasia was critically appraised. A literature search revealed a mix of single group, case series, and case-control studies with an array of statistical treatments. The evidence for this method ranged from suggestive to compelling. Experimental limitations and recommendations for clinical application are discussed.

Introduction

Injury to the brain through dementia, tumour, stroke, and other conditions affecting the language-dominant hemisphere can result in aphasia, which is an impairment that may disrupt the comprehension and/or expression of language through a variety of modalities. For instance, 21-38% of stroke patients present with aphasia in the acute stage, which can persist as a chronic condition and may negatively affect an individual's return to work, social activity, relationships, and life participation (Hosomi et al., 2009). The ability to predict a patient's language outcomes early in the acute stage would therefore be of indisputable value for the patient, their family, and their healthcare team for the purposes of discharge and rehabilitation planning. Prognostication has typically been attempted by considering a variety of known factors including lesion size, patient's age, gender, and level of education, and initial stroke and aphasia severity; however, the specificity of these measures tends to be too low for reliably predicting an individual's outcome (Hosomi et al., 2009). Furthermore, verbal assessment of patients in acute care following stroke or neurosurgery may be impeded by a patient's lack of consciousness, cooperation, or general health, which leads to delays in the examination for and characterization of aphasia (Hosomi et al., 2009). An objective anatomical measure early in the acute stage would thus be beneficial for outcome prediction.

The arcuate fasciculus (AF) is a white matter neural tract important for language function and it has been implicated in contributing to language disruption in some aphasia types (Breier et al., 2008). An association fibre tract, the AF is hypothesized to be part of the superior longitudinal fasciculus, connecting Wernicke's area in the superior temporal gyrus with Broca's area in the inferior frontal gyrus (Makris et al., 2005). Until recently, the AF could not be accurately visualized and evaluated in vivo because conventional magnetic resonance imaging (MRI) fails to differentiate the AF from adjacent neurological structures (Kim et al., 2011). A fairly new technique, diffusion tensor tractography/ imaging (DTT/DTI) takes advantage of the directional anisotropy of white matter tracts: because the movement of water through white matter tracts is much more directional than it is in gray matter, DTT can track this water mobility through intact fibres (Muthusami et al., 2014). This process permits the three-dimensional reconstruction of fibre tracts, depicting the connectivity (as well as disruptions in connectivity) between different cortical regions of the brain (Breier et al., 2008). It may therefore have potential to characterize the damaged AF in the acute stage and throughout recovery. If such a tool could be used to correlate AF injury with language outcome prediction in patients with aphasia, the utility of this anatomically-based method of prognosis would have remarkable clinical implications.

An evaluation of the existing literature would be valuable to determine whether DTT imaging of the AF is currently or has potential to be a valid anatomical method to predict language recovery in patients with aphasia.

Objectives

The primary objective of this inquiry is to conduct a critical review of the existing literature to determine the clinical utility of diffusion tensor tractography imaging of the arcuate fasciculus in predicting language outcomes for patients with aphasia.

Methods

<u>Search Strategy</u>

Peer reviewed articles pertaining to the research question were discovered using computerized databases, including: PubMed, Medline, and Google Scholar. The database search was conducted using the following keywords:

[(diffusion tensor tractography) and (aphasia) and (prognosis)]
[(diffusion tensor tractography) and (arcuate fasciculus)]

Selection Criteria
Studies selected for inclusion were required to utilize diffusion tensor tractography to image the arcuate fasciculus in patients with aphasia as these factors relate to language outcome prediction.

Data Collection
The aforementioned literature search and selection criteria yielded six appropriate studies: three studies were single group design; one was a cases series; and two were case-control studies.

Results

Single Group Design
In single group design studies, the results of one group are analyzed with no comparison/control group to weigh the results against, which inherently makes the design lack experimental power. Typically, single group design would utilize a pre- and post-test, but in some cases may be post-test only. The latter may occur in particular with retrospective studies where pre-test data was not available. Single group design may be deemed appropriate in cases where historical control groups exist.

Breier et al. (2008) conducted a single group (post-test only) study to investigate the relationship between data derived from DTT in three white matter tracts, including the AF, and language deficits in 20 patients at least one month following a left hemisphere stroke. Participant inclusion criteria was well described, albeit only broadly defined, which may have potentially confounded the author's results. Gold standard language assessments were used to evaluate three language functions (repetition, comprehension, and naming) and the DTT imaging protocol was described in detail and standardized across participants. The authors appropriately selected fractional anisotropy (FA) of the white matter tracts as their DTT parameter and the measurement of this was well described.
Within-groups MANOVA was appropriately used to examine relationships between each fibre tract's FA value and each of the three language functions. Results indicated a significant relationship between the DTT-derived FA value of the left-hemisphere AF and repetition of spoken language at least one month post-stroke. The higher the severity of white matter tract damage (evidenced by a lower AF value), the poorer the performance in repetition tasks.
Overall, this study provided suggestive evidence for the predictive value of DTT-derived data on language performance following a left hemisphere stroke.

Forkel et al. (2014) conducted a longitudinal pre- and post-test single group study investigating how lateralization of the AF (in the non-language dominant hemisphere) affects language recovery in 18 post-stroke patients with aphasia. Patient inclusion criteria were well-described and appropriate. Language was assessed using gold standard measures within 10 days and then again at 6 months of stroke onset. DTT imaging was completed on the AF within 2 weeks of stroke onset. In addition to the DTT-derived parameter of AF volume, they included a small group of known predictor variables (age, sex, and lesion size), to understand the magnitude of contribution for each variable on language prognosis. A detailed description of their calculation of AF segment volume was provided.
Appropriate hierarchical linear regression was completed to evaluate the contribution of each prognostic variable on language change across testing points (i.e., language recovery). Results revealed a significant association between language recovery and the volume of the right AF: by adding volumetric data of the right AF to the participant's age information, they could raise the predictive power of their model from 28% to 57%.
Overall, this study provided compelling evidence that lateralization to the right AF as captured by DTT is a significant predictive factor in aphasia recovery in left-hemisphere strokes.

In another pre- and post-test single group design, Hayashi et al. (2012) used DTT to study whether changes in the AF of 7 participants undergoing brain tumour resections in the language-dominant hemisphere could predict language outcomes. Participant inclusion criteria was well described. Pre- and post-operatively, language performance was assessed with a gold standard measure. DTT imaging was also conducted before and after tumour resection, the protocol of which was described in detail. The DTT parameter that the authors selected was the change in visualization of the AF, which was represented by a pre- and post-operative relative ratio (RRAF). The calculation of this was appropriate and outlined in detail.
The relationship between the RRAF and changes in language function pre- and post-operatively were analyzed by appropriately calculating a Pearson Correlation Coefficient (PCC). The results revealed a positive relationship between a RRAF (i.e., increased

visualization of the AF) and improved language performance following a tumour resection.
Overall, this study provided suggestive evidence that increased visualization of the left arcuate fasciculus post-operatively was significantly correlated with and predicted improved language function.

Case Series

Case series designs are comparable in strength to single group designs. In these studies, a series of subjects who share a condition or exposure (e.g. left hemisphere stroke) are described. Since case series lack robust statistical analyses, experimental power for generalization is limited, nor can they establish cause and effect relationships. While case series may be susceptible to selection bias and lack comparator groups, they may be justified in some cases, such as when working with rare disorders or when developing a new hypothesis as a precursor to a larger scale study.

Kim et al. (2011) conducted a case series study to examine the clinical utility of DTT for characterizing the presence and severity of left AF injury in 5 left hemisphere stroke patients to assess potential to extract prognostic information for language outcomes. Participant inclusion criteria was appropriate and well described and 7 additional age- and sex-matched controls were recruited to establish quasi-normative data for AF parameter comparison. Language performance was assessed using gold standard measures and composite and domain-specific language scores were appropriately reported. The authors chose to survey 3 different DTT parameters to assess for clinical utility: 1) the ability to fully reconstruct the left AF using DTT; 2) FA values; and 3) number of AF fibres. The DTT imaging protocol and data collection were well described and standardized across participants.
Because of the nature of case series designs, little cross-participant statistical analysis was conducted beyond comparing DTT parameters of each participant with the data collected from controls to represent uninjured AF data. This approach was appropriate, but the scant number of controls puts into question their true representativeness of a normative population. The authors concluded that they established the clinical usefulness of DTT by comparing language outcomes to left AF integrity in each case and were further able to explain language outcomes by identifying lesions in the AF using DTT that conventional MRI failed to capture.
While the evidence for DTT of the left AF being clinically useful for language outcome prognostication that this study provided was suggestive, they presented a sound argument for further investigation into this method.

Case-Control Study

The case-control study is considered a stronger experimental design than a single group or case series study. Case-control studies have the advantage of comparing an experimental group with a control group and are often done retrospectively. Participant group allocation is determined based on a particular condition or outcome of the participants and is therefore not randomized. Challenges in validating and controlling for additional variables may confound the strength of a case-control study and the study's outcomes may be affected by information bias.

Kim and Jang (2013) followed up their 2011 research by conducting a retrospective case control study to evaluate whether DTT could contribute to positively predicting language outcomes for 25 patients with aphasia secondary to left-hemisphere peri-arcuate fasciculus strokes. The participant inclusion criteria were thorough, well described, and appropriate and a nonconcurrent group of 12 age- and sex-matched controls was recruited. Gold standard language measures were employed to assess early (10-30 days post-stroke) and late (6 months post-stroke) language performance. DTT imaging was conducted within 30 days of stroke onset, the protocol of which was described in detail. The authors evaluated 3 different DTT parameters for the AF: 1) fractional anisotropy (FA); 2) apparent diffusion coefficient (ADC: a measure of the magnitude of water diffusion in neural tracts); and 3) ability to reconstruct the AF fully. The latter-most parameter was used to divide the 25 participants into three groups, reflecting the severity of injury (and integrity of) the AF.
A Pearson Correlation Coefficient analysis was used appropriately to evaluate whether relationships existed between each of the DTT parameters and the participants' language scores over the two time points (i.e., their language recovery). The author's findings revealed that the participant groups whose left AF could at least be partially reconstructed on DTT within thirty days of stroke had significantly improved language outcomes over the participants whose left AF could not be reconstructed (i.e., more severe injury). Further, the FA values of those participants whose left AF could at least be partially reconstructed were positively correlated with improved aphasia outcomes at 6 months post-stroke, evidenced by both their composite language score and repetition score.
This study provided compelling evidence that language performance in patients with aphasia was positively correlated with the ability to reconstruct the left AF using DTT.

Hosomi et al. (2009) conducted a retrospective case-control study to evaluate whether information derived from DTT could positively predict language recovery in

13 patients with left MCA strokes at discharge. Participant inclusion criteria was well described and a nonconcurrent age- and sex-matched control group of 10 individuals was recruited. Participants were divided into 2 groups depending on whether they did or did not have aphasia at discharge, which was determined with a bedside scale of stroke severity that included a language screen. The procedures for the DTT imaging and data collection were well described and were all conducted within 2 days of stroke onset. The authors appropriately selected 2 DTT parameters for their inquiry: indices based on the asymmetry of the left vs. right AF for each participant using 1) fractional anisotropy (FA) and 2) the number of AF fibres. The computation of these was described in detail.

Appropriate implementation of two non-parametric statistical analyses was completed to first assess intraindividual differences between left and right AFs and then to evaluate intergroup differences between participants and controls. Results revealed that in controls and the nonaphasia group, the number of AF fibres was significantly higher in the left AF than in the right, while the aphasia group did not demonstrate the same pattern of leftward asymmetry. Further, this loss of leftward asymmetry showed high sensitivity (.83) and specificity (.86) for predicting aphasia at discharge.

Overall, this study provided compelling evidence that the DTT-derived AF fibre number asymmetry index positively predicted the presence of aphasia in left MCA stroke patients at discharge.

Discussion

Of the 6 studies reviewed, half provided suggestive (Breier et al. 2008; Hayashi et al., 2012; Kim et al., 2011) and half provided compelling (Forkel et al., 2014; Hosomi et al., 2009; Kim & Jang, 2013) early evidence for using DTT imaging of the arcuate fasciculus to predict language outcomes in patients with aphasia. Though the method shows promise, clear limitations emerged through this review, which are pertinent to consider before full clinical adoption.

DTT Parameters

For appropriate clinical application, DTT parameter selection should be appropriate, valid, and standardized. There were a variety of measures extracted from DTT of the arcuate fasciculus across the studies including diffusion magnitude, fractional anisotropy (FA), number of AF fibres, volumetric data, and AF reconstruction. Further, inter-study variation between the predictive value of the AF of the language dominant vs. non-dominant hemispheres also varied. It was also evident that parameters varied in sensitivity depending on the context (e.g. location of lesion). For instance, FA is a measure of directionality of water through association fibres and therefore represents their structural integrity and alignment (Muthusami et al., 2014). Thus, the face validity of using FA as a DTT parameter to measure fibre integrity is high, as in the studies of Breier et al. (2008) and Kim and Jang (2013). Forkel et al. (2014) pointed out, however, that volumetric data of the left and right AF derived from DTT represents two different things. The right AF volume represents an intact neural tract that has been unaffected by a lesion. However, the left AF volume measure represents a neural tract that has been injured and potentially disrupted. Thus, it may be possible that a volumetric parameter is more aptly measured in the non-dominant hemisphere, while another parameter, such as fractional anisotropy, may be more revealing for outcome prediction in the language dominant hemisphere.

Hosomi et al. (2009) took a unique approach in choosing a DTT parameter by calculating asymmetry indices for FA and the number of AF fibres. Because they did not have the luxury of pre- and post-tests or repeated DTT measures due to the limitations of their case-control study design, using each patient's right AF measures to control for individual differences captured the condition of the participants' AF at discharge as validly as possible. In utilizing this asymmetry index, Hosomi et al. (2009) were able to correct for one of the confounds of Forkel's study observations, which was that by looking at the left AF alone, volumetric measures of the AF are not as meaningful. But the use of the right AF as a control to generate an index corrects for this design issue.

Spontaneous vs. Therapy-Mediated Language Recovery

Of the 6 studies evaluated, only that of Hosomi et al. (2009) reported definitively that their participants received speech and language therapy during their hospital stay. The rest of the studies either did not report it or did not have that data available to them because of their retrospective nature. As such, it is difficult to assess the role that language therapy played in prognostication and whether the outcome prediction data derived from DTT represented spontaneous or therapy-mediated language recovery. To be sure, it would be unethical to design an experiment where a group of patients is denied language therapy, however controlling for the participation in language therapy could be a consideration for future research.

Timing of DTT

A clear limitation of the retrospective research designs included in this review was the inability to control for specific variables related to timing of DTT imaging. For instance, all of the participants in Hosomi et al.'s 2009 study were imaged within 2 days of stroke onset, while Forkel et al. (2014) conducted their initial DTT pre-test in the subacute stage, and Kim et al. (2011) reported a time range of 13-106 days between stroke onset and DTT. Without controlling for timing of the DTT imaging among subjects between stroke onset and imaging, post stroke reorganization could be an experimental confound and limit the ability to establish relationships between experimental variables.

Location and Extent of Lesion

Further, the location and extent of neurological involvement of the lesion is a significant factor that must be considered. While Kim and Jang (2013) limited their participants to those who had peri-arcuate fasciculus strokes, most of the research included in this review demonstrated a range and severity of neurological structure implication. The current review looked only at the arcuate fasciculus, but there may be other white matter tracts more aptly evaluated, depending on the extent and location of the lesion. Also, Hayashi et al. (2012) pointed out that when imaging brain tumours of varying sizes and locations, the compressional effects of the tumour on the surrounding anatomy make the task of identifying the DTT regions of interest more subjective and difficult to duplicate.

Clinical Implications

A review of the body of evidence for using DTT imaging of the arcuate fasciculus as an anatomical predictor of language outcomes in patients with aphasia indicates early promise for this method. Despite limitations inherent in the included research designs, evidence ranged from suggestive to compelling for pursuing clinical implementation.

Before putting DTT-derived data of the arcuate fasciculus to clinical use, effort must be made to select and standardize the most valid DTT parameter(s) for inclusion, as well as the timing and protocol of imaging. A clearer understanding for the role of language therapy in this prognostic model must also be achieved and would be a priority for future inquiry. Further, the population of individuals with aphasia is heterogenous with a range of factors to consider including, but not limited to: etiology, neurological structure involvement, and severity of injury. Before clinical implementation, further investigation needs to be conducted to control for some of these factors. For instance, DTT imaging of the arcuate fasciculus may not be appropriate in patients whose stroke lesions demonstrate no involvement of this area. Currently, DTT imaging of the AF may be best utilized as part of a broader set of prognostic indicators rather than as a sole predictor of language recovery in patients with aphasia.

References

Breier, J.I., Hasan, K.M., Zhang, W., Men, D., & Papanicolaou, A.C. (2008). Language dysfunction after stroke and damage to white matter evaluated using diffusion tensor imaging. *American Journal of Neuroradiology*, 29: 483-487

Forkel, S.J., Thiebaut de Schotten, M., Dell’Acqua, F., Kalra, L., Murphy, D.G.M., Williams, S.C.R., & Catani, M. (2014). Anatomical predictors of aphasia recovery: a tractography study of bilateral perisylvian language networks. *Brain: A Journal of Neurology*, 137: 2027-2039

Hayashi, Y., Kinoshita, M., Nakada, M., & Hamada, J-I. (2012). Correlation between language function and the left arcuate fasciculus detected by diffusion tensor imaging tractography after brain tumor surgery. *Journal of Neurosurgery*, 117: 839 -843

Hosomi, A., Nagakane, Y., Yamada, K., Kuriyama, N., Mizuno, T., Nishimura, T., & Nakagawa, M. (2009). Assessment of arcuate fasciculus with diffusion-tensor tractography may predict the prognosis of aphasia in patients with left middle cerebral artery infarcts. *Neuroradiology*, 51: 549-555

Kim, S.H., Lee, D.G., You, H., Son, S.M., Cho, Y.W., Chang, M.C., Lee, J., & Jang, S.H. (2011). The clinical application of the arcuate fasciculus for stroke patients with aphasia: A diffusion tensor tractography study. *Neurorehabilitation*, 29: 305-310

Kim, S.H., & Jang, S.H. (2013). Prediction of aphasia outcome using diffusion tensor tractography for arcuate fasciculus in stroke. *American Journal of Neuroradiology*, 34: 85-90

Makris, N., Kennedy, D.N., McInerney, S., Sorensen, A.G., Wang, R., Caviness, V.S. Jr., Pandya, D.N. (2005). Segmentation of subcomponents within the superior longitudinal fasciculus in humans: a quantitative, in vivo, DT-MRI study. *Cerebral Cortex*, 15(6): 854-869

Muthusami, P., James, J., Thomas, B., Kapilamoorthy, T.R., & Kesavadas, C. (2014). Diffusion tensor imaging and tractography of the human language pathways: Moving into the clinical realm. *Journal of Magnetic Resonance Imaging*, 40: 1041-1053