North-American Lidcombe Program file audit: Replication and meta-analysis

SARITA Koushik1, SALLY HEWAT1, ROSALEE C. SHENKER2, MARK JONES3, & MARK ONSLOW4

1The University of Newcastle, Australia, 2The Montreal Fluency Centre, Quebec, Canada, 3The University of Queensland, Brisbane, Australia, and 4Australian Stuttering Research Centre, The University of Sydney, Australia

Thousands of North American clinicians have trained for the Lidcombe Program of Early Stuttering Intervention, yet there are no benchmark data for that continent. This retrospective file audit includes logistical regression of variables from files of 134 children younger than 6 years who completed Stage 1 of the Lidcombe Program. Benchmarking data for clinic visits to Stage 2 is available for these files. Meta-analysis supplements worldwide Lidcombe Program benchmark data. The median number of clinic visits to Stage 2 was 11. High pre-treatment stuttering severity predicted more clinic visits than low severity. A trend toward statistical significance was found for the frequency of clinic visits. Frequent attendance of mean less than 11 days was associated with longer treatment times than infrequent attendance of mean 11 days or more. Results for North America were consistent with benchmark data from the UK and Australia. The mean attendance trend is clinically important and requires further investigation because of its potential clinical significance.

Keywords: Stuttering, pre-school, Lidcombe Program.

Introduction

The onset of stuttering typically occurs during the pre-school years. A recent community cohort ascertained prior to stuttering onset, and confirmed by expert diagnosis, reported 3-year cumulative stuttering incidence at 8.5% with median onset age of 29.9 months (Reilly, Onslow, Packman, Wake, Bavin, Prior, et al., 2009). This value is higher than the previously reported incidence of ~ 5% (Andrews & Harris, 1964; Mansson, 2000). Some children recover from stuttering without treatment, with reports of natural recovery rate within 4 years of onset at 74% (Bloodstein, 1995; Yairi & Ambrose, 1999). If untreated, stuttering can lead to negative consequences. Negative peer responses occur during the pre-school and school years (Davis, Howell, & Cooke, 2002; Langevin, Packman, & Onslow, 2009). Persistent stuttering can lead to social phobia, maladjustment, and anxiety disorders (Craig & Calver, 1991; Hayhow, Cray, & Enderby, 2002). Stuttering adults are at extreme risk of social phobia (Iverach, O’Brien, Jones, Block, Lincoln, Harrison, et al., 2010). Anxiety-related disorders impede efficacious speech treatment (Iverach, Jones, O’Brian, Block, Lincoln, Harrison, et al., 2009a). Therefore, an efficacious early intervention program is essential to assist clinical decision-making shortly after onset to avoid these consequences.

Randomized controlled trials (RCTs) are a research design that is commonly used to provide evidence that a treatment is efficacious. The advantage of RCTs is that any differences found between groups are likely due to the intervention, thereby minimizing bias (Hoffman, Bennett, & Del Mar, 2010). Presently, the Lidcombe Program is the only treatment for pre-school-age children that has randomized controlled clinical trials evidence (Jones, Onslow, Packman, Williams, Ormond, Schwarz, et al., 2005; Lewis, Packman, Onslow, Simpson, & Jones, 2008). A recent meta-analysis of randomized clinical evidence showed an odds ratio of 7.5 for 130 treated children (Onslow, Jones, Menzies, O’Brien, & Packman, 2008). In other words, children who were treated with the Lidcombe Program had 7.5-times the odds of attaining minimal stuttering compared to controls.

Correspondence: Sarita Koushik, Humanities and Social Science, Faculty of Education and Arts, The University of Newcastle, Callaghan, NSW 2308, Australia. Tel: 61-2-4921 6414. E-mail: sarita.koushik@gmail.com

ISSN 1754-9507 print/ISSN 1754-9515 online © The Speech Pathology Association of Australia Limited

Published by Informa UK, Ltd.

DOI: 10.1080/17549507.2011.538434
The Lidcombe Program (Onslow, Packman, & Harrison, 2003) is a behavioural treatment based on verbal response contingent stimulation. Stage 1 involves weekly parent and child clinic visits. This stage continues until the child reaches zero or near-zero levels of stuttering for three consecutive clinic visits, which are the criteria for entering Stage 2. The goal during Stage 2 is maintenance of zero or near-zero stuttering for at least 1 year.

In two previous studies, case variables from clinical files were examined for two specialized stuttering clinics in Australia and the UK. The purposes were to determine if the duration of treatment with the Lidcombe Program could be predicted. For those clinics, large, independent cohort retrospective file audits were conducted for children who completed Stage 1 of the program (Jones, Onslow, Harrison, & Packman, 2000; Kingston, Huber, Onslow, Jones, & Packman, 2003). The clinical files of 250 Australian and 66 British children were independently audited. Pre-treatment stuttering severity, measured with percentage of syllables stuttered (%SS), was found to be a significant predictor in both studies. Children with more severe stuttering (5.0% SS or more) required more clinic visits to reach Stage 2 than children with less severe stuttering (less than 5.0% SS). Both studies reported the median number of clinic visits to Stage 2 to be 11.

The data from the Australian and British studies were pooled (n = 316) for meta-analysis (Kingston et al., 2003). Meta-analysis increased the sample size, thus the statistical power. Children whose pre-treatment stuttering was more severe were 3.5-times as likely (p < 0.0001) to require more clinic visits to complete Stage 1 than less severe children. Further, children who had been stuttering less than 12 months had twice the odds (p = .013) of requiring more clinic visits to complete Stage 1 than children stuttering for more than 12 months. The meta-analysis improved benchmarking data for clinicians who use the Lidcombe Program.

Clinical benchmarking contributes knowledge about the process of clinical care and outcomes (Higgins, 1997). In speech-language pathology there is little published on clinical benchmarking (Hunt & Slater, 1999) and this is so for stuttering (Yaruss, LaSalle, & Conture, 1998). By providing benchmarking data, healthcare professionals can compare treatment delivery to the standard. This allows management of health services and allocation of funds for treatment approaches.

The Lidcombe Program Training Consortium is an international group, with members in eight countries, who provide Lidcombe Program training worldwide in English and non-English speaking countries (Australian Stuttering Research Centre, 2009). In North America since 2001, more than 3000 clinicians in the US and Canada have received a 2-day basic Lidcombe Program skills workshop. However, large cohort retrospective recovery studies have not been conducted for that continent. The purposes of this study are to replicate the Jones et al. (2000) benchmark study in North America and to combine data with Australian and British studies for meta-analysis for the benefit of all clinical communities.

**Method**

The Jones et al. (2000) procedures were replicated and extended. A retrospective file audit included four clinical sites from the US and one site from Canada. Personnel from the clinics extracted and de-identified requisite information from files of all children who were treated with the Lidcombe Program. Fifteen clinicians with varying levels of experience treated the children with the procedures as described in the manual (Packman, Webber, Harrison, & Onslow, 2008). All treating clinicians had received a 2-day Lidcombe Program basic skills workshop.

File data were collected from 165 children who attended the clinics during the years 2002-2009 and began treatment when younger than 6 years. The numbers of clinical files contributed from each clinic were 54, 50, 31, 20, and 10. Children were only included in the analyses, if they reached Stage 2, in order to provide clinical benchmarks for the children who completed the first stage of the program. The criteria for Stage 2 entry in the Packman et al. (2008) manual are “(1) %SS less than 1.0 within the clinic, and (2) Severity rating (SR) scores for the previous week of 1 or 2, with at least four of these being 1” (Australian Stuttering Research Centre, 2008, p. 8). These criteria need to be achieved for three consecutive clinic visits.

Non-progress to Stage 2 occurred in 27 cases (13.5%), therefore these files were withdrawn from further analysis. Reasons for non-progress occurred in 10 cases because parent schedules conflicted with available clinic times, in four cases because the children lost funding and could not continue on a private pay basis, in two cases because of concurrent pressing medical treatment, in five cases because the families felt that progress with the Lidcombe Program was slower than expected, and for six files no reason was reported. Of the 27 children who did not progress to Stage 2, according to file data, 20 decreased their stuttering severity by more than 2.0%SS from pre-treatment until the time of drop-out, three children showed no change, and four children had missing file data at the time of drop-out.

The remaining 138 children, 105 boys and 33 girls, progressed to Stage 2 of the Lidcombe Program. Therefore, the number of children from each clinic included in the analysis was 46, 41, 27, 20, and 4, respectively. Data from one clinic were removed because only four children reached Stage 2, which therefore would not make a meaningful contribution to the analyses. The final analyses are based on 134 children.
**Variables**

**Dependent variable**

The dependent variable was the number of clinic visits required for entry to Stage 2 of the Lidcombe Program. This variable was categorized to represent short and long treatment duration. Short treatments were defined as fewer than 12 clinic visits and long treatments were 12 visits or more. Categorizing the dependent variable was decided upon because treatment time as a continuous variable did not meet the requisite assumptions for least squares regression.

**Predictor variables**

The following four predictor variables used by Jones et al. (2000) and Kingston et al. (2003) were obtained from each clinical file: gender, age at the first treatment visit, onset-to-treatment interval, and stuttering severity (%SS) at the first treatment visit. The variables were categorized identically to Jones et al. (2000). Categorization avoided the assumption that any relationship with the dependent variable would be linear. Age at the first treatment visit was categorized into younger than 4 years or 4 years and older. Stuttering severity at the first treatment visit was categorized as less severe being below 5.0%SS and more severe being greater or equal to 5.0%SS. Onset to treatment interval—the time between reported stuttering onset and the first treatment visit—was categorized to shorter than 12 months and longer than 12 months. That categorization reduced reliance on parent recall of onset.

Although the Lidcombe Program manual specifies that treatment be provided with weekly clinic visits during Stage 1 (Packman et al., 2008), there are many reasons beyond a clinician’s control why this may not occur. Failures to attend clinic appointments occur for various reasons, some of which include illness, scheduling conflicts, or vacations. Additionally, two reports suggest that clinicians deviate from the weekly visit requirement in order to manage caseloads (O’Brien, Iverach, Jones, Onslow, Packman, & Menzies, 2009; Rousseau, Packman, Onslow, Dredge, & Harrison, 2002). Therefore, mean days between clinic visits was calculated for each child and categorized into frequent visits (fewer than 11 visits) and infrequent visits (11 visits or more). This categorization was based on the mean value calculated for days between clinic visits for the cohort.

**Results**

Analyses used SAS for Windows, version 9.2 (SAS Institute, Cary, NC). Goodness-of-Fit statistics were used to assess the final logistic models. Descriptive statistics for the predictor variables are presented in Table I. The median age at the first treatment visit was 4.1 years (SD = 9.5), median onset-to-treatment interval was 13 months (SD = 10.2), median days between clinic visits was 10 (SD = 5.8), and median %SS at the first treatment visit was 5.0%SS (SD = 5.1). The median %SS at the first treatment visit was calculated for 131 clinic files because this datum was missing from three files.

**Median clinic visits by clinic site**

Evidence of heterogeneity was found between the clinic sites (log-rank p = .01), therefore the presented analyses are stratified by clinic. A Kaplan-Meier survival analysis is a descriptive statistical procedure for the time to event variables (Kaplan & Meier, 1958). It is used in cases where time is the most prominent variable and involves the generation of survival plots. For the North American data, a survival analysis is used as a tool to measure the required time (i.e. measured in number of clinic visits) to reach Stage 2 (i.e. event). Figure 1 represents the cumulative proportion of children who attained Stage 2 by the number of clinic visits, stratified by clinic site. The median clinic visits is represented by .50 of the proportion of children reaching Stage 2 or, in other words, where 50% of all children reached near-zero stuttering. For all clinics, the median clinic visits to attain Stage 2 were similar except for one clinic. Medians for the clinics were 11, 10, and 14, but 23 visits for one clinic. Clearly, the outlying data for one clinic requires further exploration. For the purposes of providing benchmarking data for number of clinic visits, the data from this clinic was included in the group analysis.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Age at first treatment visit (months)</th>
<th>Onset to treatment interval (months)</th>
<th>%SS at first treatment visit</th>
<th>Days between clinic visits</th>
<th>Number of clinic visits to Stage 2</th>
<th>Number of clinic visits to Stage 2 (not including outlier)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NValid</td>
<td>124</td>
<td>122</td>
<td>131</td>
<td>132</td>
<td>134</td>
<td>114</td>
</tr>
<tr>
<td>Missing</td>
<td>10</td>
<td>12</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>Mean</td>
<td>49.6</td>
<td>15.9</td>
<td>6.3</td>
<td>11</td>
<td>14.1</td>
<td>12.4</td>
</tr>
<tr>
<td>Median</td>
<td>49.5</td>
<td>13</td>
<td>5</td>
<td>10</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Mode</td>
<td>51</td>
<td>8</td>
<td>3</td>
<td>10.5</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>9.5</td>
<td>10.2</td>
<td>5.1</td>
<td>5.8</td>
<td>7.5</td>
<td>5.8</td>
</tr>
<tr>
<td>Range</td>
<td>31–71</td>
<td>1–53</td>
<td>0.3–32</td>
<td>2.3–49.9</td>
<td>4–44</td>
<td>4–44</td>
</tr>
</tbody>
</table>
**Median clinic visits for the group**

A Kaplan-Meier plot for the number of clinic visits is presented in Figure 2. For the 134 children, the median number of clinic visits was 12. The 90th percentile was 22 visits. If the files from the outlying clinic were not included in the analysis, the median and 90th percentile value decreased to 11 and 21, respectively. Four children were below 1.0%SS at the first clinic visit. To confirm that these children did not affect the median value for the whole cohort, the data were reanalysed without them. For the 130 files, the median number of clinic visits to Stage 2 remained at 12.

**Logistic regression**

To determine the relationship between the dependent variable and all five predictor variables, a univariable logistic regression analysis was performed. For predictor variables, one category was specified as the reference and the non-reference value was measured for significance. The odds ratio is a measure of the strength of relationship between two variables. If the odds ratio for the non-reference value is 1.0 there is no difference between the groups. Table II shows the results of the univariable regression, stratified by clinic site.

The data showed no evidence of an association between number of clinic sessions and age, gender, or onset-to-treatment interval. However, there was strong evidence that higher severity is associated with more clinic visits (p = .004). Children with stuttering severity of 5.0%SS or more had approximately a 4-fold increased odds of requiring 12 or more visits than the milder group. Additionally, there is some evidence that frequent clinic attendance is associated with more clinic visits to Stage 2 (p = .04). Children who attended the clinic more often than every 11 days had more than twice the odds of requiring longer than 12 clinic sessions compared to children who attended the clinic infrequently.

A multivariable logistic regression analysis showed similar results to the univariable analysis. The association between frequency of attendance and number of clinic sessions approaches statistical significance (odds ratio = .47, p = .07). For the variable severity of stuttering, the association between severity and number of clinic sessions was almost unchanged (odds ratio = 3.7, p = .01).

**Goodness-of-fit**

Goodness-of-fit statistics were used to assess the final logistic models. The c-statistic indicates how well the model distinguishes between children taking a shorter and children taking a longer number of clinic sessions, where 0.5 indicates a model that is not predictive and 1.0 indicates a model that predicts perfectly. Pearson’s chi-square test was used to test that the models did not provide a poor fit to the data.

Results for the North American data are presented in Table III. The final logistic model had a c-statistic

<table>
<thead>
<tr>
<th>Variable</th>
<th>Odds Ratio</th>
<th>95% Confidence Interval</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>%SS at first clinic visit</td>
<td>1.0*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5%SS+</td>
<td>3.8</td>
<td>1.5 – 9.4</td>
<td>.004</td>
</tr>
<tr>
<td>Onset-to-treatment interval</td>
<td>1.0*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 mths+</td>
<td>1.0*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>1.0*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1.0*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.84</td>
<td>0.34 – 2.1</td>
<td>.07</td>
</tr>
<tr>
<td>Age</td>
<td>1.0*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;4 years+</td>
<td>1.0*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attendance</td>
<td>1.0*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>More frequent</td>
<td>1.0*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less frequent</td>
<td>0.44</td>
<td>0.20 – 0.96</td>
<td>.04</td>
</tr>
</tbody>
</table>

*reference category.

---

**Figure 1.** Kaplan-Meier plot of cumulative proportion of participants who attained Stage 2 by clinic site

**Figure 2.** Kaplan-Meier plot of cumulative proportion of 134 participants who attained Stage 2 by number of clinic visits

---

S. Koushik et al.
Table III. Results of the univariable logistic regression (Australian, British and North American cohorts)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Odds Ratio</th>
<th>95% Confidence Interval</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>%SS at first clinic visit</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 5 %SS</td>
<td>1.0*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5–9.9 %SS</td>
<td>2.3</td>
<td>1.4 – 3.7</td>
<td>0.0008</td>
</tr>
<tr>
<td>10 %SS+</td>
<td>5.2</td>
<td>2.5 – 10.6</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Onset-to-treatment interval</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 12 mths</td>
<td>1.0*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 mths+</td>
<td>0.76</td>
<td>0.50 – 1.1</td>
<td>0.18</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1.0*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.70</td>
<td>0.44 – 1.1</td>
<td>0.14</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 4 years</td>
<td>1.0*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 years+</td>
<td>0.87</td>
<td>0.59 – 1.3</td>
<td>0.49</td>
</tr>
</tbody>
</table>

*reference category.

of .79 and Pearson’s chi-square = 10.4, df = 10, p = .4. These statistics indicate no evidence of lack of fit and a model that can reasonably distinguish between children taking shorter and children taking longer numbers of clinic sessions.

**Meta-analysis**

The data collection methods for the present study were identical to those of Jones et al. (2000) and Kingston et al. (2003). However, those studies did not collect data for number of days between clinic visits; therefore that variable was not included in the meta-analysis. The data sets for 444 children who attained Stage 2 of the Lidcombe Program for the three studies were combined. For the purpose of the meta-analysis, severity was re-categorized into three levels. The categories of mild (0–4.9%SS), moderate (5.0–9.9%SS), and severe (10.0%SS+) were used to report findings. Results are presented in Table III.

Based on meta-analysis of the 444 cases, there was no evidence of a correlation between age, gender, onset-to-treatment interval, and treatment duration. However, there was strong evidence of correlation between stuttering severity and treatment duration. Based on the Kaplan-Meier analysis and log-rank test there was strong evidence that increasing severity is associated with increased number of clinic visits (p < .0001). For the group, the median number of clinic visits to Stage 2 was 11. Based on the re-categorization into three levels, the median number of clinic visits to Stage 2 was 10 for mild, 12 for moderate, and 14 for severe pre-treatment severity of stuttering. More specifically there was strong evidence that children with moderate pre-treatment severity had more than double the odds of a longer duration of treatment (p = .0008) compared to children with mild pre-treatment severity of stuttering.

The Goodness-of-Fit final logistical model for the meta-analysis had a c-statistic of .67 and Pearson’s chi-square = 13.4, df = 12, p = .34. These statistics indicate no evidence of lack of fit and a model that has some ability to distinguish between children taking shorter and children taking longer numbers of clinic sessions; however, a limitation is that frequency of attendance was not able to be included in the model.

**Discussion**

The present study replicated the Jones et al. (2000) study and combined the data sets with the Australian and British studies for meta-analysis to provide benchmarking data. Of the five participating North American clinics, one was removed from the analysis because of the small number of contributed files. The median visits to complete Stage 1 by clinic site were similar for three clinics at 11, 10, and 14. However, one clinic reported a median of 23 visits. The difference for this clinic could not be explained by higher severity of stuttering. Possible explanations could be due to differences in the service delivery of the Lidcombe Program, differences in clinician experience, or data errors. To provide benchmarking data for clinic visits to Stage 2, the files from this clinic were included in the final analysis. Pre-treatment stuttering severity was found to be a significant predictor of treatment time for the North American cohort. Stuttering severity of 5.0%SS or higher required more sessions to complete Stage 1 than lower pre-treatment severity. However, the three predictor variables onset-to-treatment interval, gender, and age at first treatment visit were not found to be significant predictors of treatment time. The variable frequency of clinic visits produced an unexpected finding. On average, children who attended the clinic frequently (average less than 11 days) required more clinic visits to complete Stage 1 than infrequent attendees (average greater than 11 days). A statistical trend in the multivariable regression showed some evidence of an association (p = .07), although marginally so. This is a clinically important trend in the data that requires further investigation because of its potential clinical significance. If further evidence were to produce similar results, this would have implications for service delivery options with the Lidcombe Program.

The number of clinic visits to Stage 2 were similar to the Jones et al. (2000) and Kingston et al. (2003) studies. In these studies, the reported median value of 11 visits was similar to the North American cohort which reported 12 clinic visits. Further, the 90th percentile value for the North American cohort was 22 clinic visits. In comparison, 90% and 95% of the Australian and British cohorts completed Stage 1 in 22 and 21 clinic visits, respectively. All three studies independently agreed on the median and 90th
95th percentile values. For treatment time to Stage 2, the Australian and British studies did not find significant results for gender and age at first treatment visit. These results were similar to those obtained from the North American cohort. These findings provide important benchmarking data for number of clinic visits. Clinicians and health care managers can utilize this benchmark information when planning and delivering early stuttering intervention services.

The meta-analysis of the three studies increased the statistical power, thus providing important benchmarking information. A highly significant predictor for treatment time was pre-treatment severity which was re-classified into three categories. The median number of clinic visits for severity was 10 for mild, 12 for moderate, and 14 for severe pre-treatment stuttering severity. The meta-analysis in Kingston et al. (2003) showed a significant correlation between onset-to-treatment interval and treatment time. However, when the North American data were included, this correlation became non-significant. A test for interaction was used to determine whether there was a differential effect in the Kingston et al. (2003) data compared to the North American data. However, this was not found to be the case (Wald Chi-Square 1.25, df = 1, p = .26).

An important finding was the agreement of median values obtained from the North American cohort independent of the Australian and British cohorts. Further, the meta-analysis provided important benchmarking data for clinical translation. It is important to note that the Australian, British, and North American studies were performed primarily in English speaking nations. Investigations of the Lidcombe Program with non-English speaking countries is required to determine if these benchmarks are achievable for countries with different languages and cultures.

Acknowledgements

This research was supported in part by Program Grant 402763 from the National Health and Medical Research Council of Australia. The authors would like to acknowledge the contribution of clinical file data from the five participating clinic sites in North America.

References


