Papers

Randomised controlled trial of community based speech and language therapy in preschool children

Margaret Glogowska, Sue Roulstone, Pam Enderby, Tim J Peters

Abstract

Objective To compare routine speech and language therapy in preschool children with delayed speech and language against 12 months of "watchful waiting."
Design Pragmatic randomised controlled trial.
Setting 16 community clinics in Bristol.
Participants 159 preschool children with appreciable speech or language difficulties who fulfilled criteria for admission to speech and language therapy.
Main outcome measures Four quantitative measures of speech and language, assessed at 6 and 12 months; a binary variable indicating improvement, by 12 months, on the trial entry criterion.

Results Improvement in auditory comprehension was significant in favour of therapy (adjusted difference in means 4.1, 95% confidence interval 0.5 to 7.6; P = 0.025). No significant differences were observed for expressive language (1.4, -2.1 to 4.8; P = 0.44); phonology error rate (-4.4, -12.0 to 3.3; P = 0.26); language development (0.1, -0.4 to 0.6; P = 0.73); or improvement on entry criterion (odds ratio 1.3, 0.67 to 2.4; P = 0.46). At the end of the trial, 70% of all children still had substantial speech and language deficits.

Conclusions This study provides little evidence for the effectiveness of speech and language therapy compared with watchful waiting over 12 months. Providers of speech and language therapy should reconsider the appropriateness, timing, nature, and intensity of such therapy in preschool children. Continued research into more specific provision to subgroups of children is also needed to identify better treatment methods. The lack of resolution of difficulties for most of the children suggests that further research is needed to identify effective ways of helping this population of children.

Introduction

Of the impairments presenting in early childhood, speech or language delay may be the most common.¹ At any one time a fifth of parents in Britain are concerned about their young child's language development.² Although there has been a shift to providing early intervention for these children, this has not been based on research evidence. Yet provision of therapy to children is estimated to consume 70% of the NHS

budget for speech and language therapy in the United Kingdom. $^{\circ}$

A systematic review has shown short term efficacy of speech and language therapy for young children in an experimental environment.³ No clear evidence exists, however, on the long term effectiveness of therapy in the context of service provision or on the natural course of early speech and language delays. In particular, the longer term course of early difficulties seems to vary for different groups of children. Some studies have suggested that 40% to 60% of children with only expressive language delay outgrow their difficulties^{4–5}; others have shown that those with a range of language problems have more persistent linguistic, literacy, and social difficulties.^{6–8}

We investigated in a pragmatic randomised controlled trial the effectiveness of speech and language therapy for preschool children as delivered in community clinics. Parallel studies involving in-depth interviews with parents and questionnaires to investigate parents' views are reported elsewhere.^{9 10}

Subjects and methods

Subject selection and baseline assessment

Children eligible for the trial were identified by 21 speech and language therapists working in 16 NHS community clinics. We considered for inclusion all children presenting to these clinics from primary care from December 1995 to March 1998. Box 1 shows the selection criteria; box 2 shows the three clinical criteria. Local research ethics committees for the three participating healthcare trusts gave approval for the trial; informed parental consent was obtained by therapists. Baseline assessments were performed before randomisation and included the preschool language scale¹¹ to determine auditory comprehension and expressive language scores (age standardised to a mean of 100 and SD of 15); phonological analysis of 22 words, yielding a percentage error rate, adapted from Pagel Paden et al¹²; a 30 minute audiotaped sample of the child's spontaneous verbal output scored from 0 to 10 on the Bristol language development scales¹³; and the daily living, socialisation, and motor skills domains of the Vineland adaptive behaviour scales (standardised as for the preschool language scale).¹⁴ Other measures of overall functioning were attention levels (range 1-6),¹⁵ symbolic play (range 1-5),¹⁶ and the therapy outcome measures tool (range 0-5).¹⁷ We also collected Research Unit, Frenchay Hospital, Bristol BS16 1LE Margaret Glogowska *research assistant* Sue Roulstone *clinical research director*

Editorial by Law

and Conti-Ramsden

Centre for Ageing and Rehabilitation Studies, School of Health and Related Studies, University of Sheffield, Northern General Hospital, Sheffield S5 7AU Pam Enderby *professor of community rehabilitation*

Department of Social Medicine, University of Bristol, Canynge Hall, Bristol BS8 2PR Tim J Peters reader in medical statistics

Correspondence to: S Roulstone sue@speech-therapy. org.uk

BMJ 2000;321:1-5

Box 1–General selection criteria

- Newly referred singleton children acquiring English in a monolingual home
- Aged under $3\frac{1}{2}$ years at initial attendance for
- speech and language therapy assessment
- · No diagnosis of severe learning difficulties or autism
- No oromotor deficits
- No primary diagnosis of dysfluency (stammering) or dysphonia (voice disorders)
- No siblings currently receiving speech and language therapy
- Children had to satisfy one of the clinical criteria (box 2)
- Be considered to have significant clinical difficulties by the speech and language therapist
- A "carer" had to attend sessions
- · Parents had to give consent

data from parents on their own education, employment, and age; family composition; child care arrangements; familial speech, language, or learning difficulties; and the child's medical, hearing, and communication history.¹⁸

Assignment

Eligible children were randomised to receive therapy or to "watchful waiting." Randomisation was stratified by the 16 clinics and by the three clinical criteria (general language, expressive language, and phonology). Although randomisation within each stratum was in blocks of six, the large number of strata meant that equal numbers in the two arms were not guaranteed. The sequence of random numbers was generated before the trial independently of the therapists. The allocation was implemented by the therapists opening sealed opaque envelopes (coloured according to the three clinical criteria) in the presence of the parents. An audit using the sequential serial numbers assigned to the children confirmed that the therapists had adhered to this procedure.

All parents were given information at baseline about the nature and severity of their child's difficulties, together with general advice. Follow up assessments at six months provided interim monitoring for both groups. Children randomised to the therapy group received the one-to-one speech and language therapy routinely offered by the therapist, starting within a

Box 2-Clinical criteria

General language group—a standardised score < 1.2 SD (standard deviation) below the mean on the auditory comprehension part of the preschool language scale¹¹

Expressive language group—a standardised score >1.2 SD below the mean on auditory comprehension but <1.2 SD below the mean on the expressive language part of the preschool language scale

Phonology group—auditory comprehension and expressive language scores > 1.2 SD below the mean but with an error rate of at least 40% in the production of fricative consonants (for example, f and s) and/or velar consonants (for example, "hard" c, "hard" g, and ng) and/or sounds occurring after a vowel among the 22 words included in the phonological analysis month of randomisation and tailored to their individual needs. Parents of children in the watchful waiting group could request therapy at any time if they were concerned about their child's progress. All children in the study were reassessed by the research therapists after 12 months; if a child in the watchful waiting group was still experiencing difficulties, two research therapists (SR and MG) provided up to 12 therapy sessions.

Blinding

The same assessments were used at 6 and 12 months as at baseline, with the exception of the Bristol language development scales and the Vineland scales, which, for lack of time, were measured only at 12 months. Assessors were blind to previous results, and every attempt was made to maintain blindness in terms of allocation. The presence of the parents meant that this was often inevitably compromised, but each child was seen by a different therapist for the two follow ups, and the language sample for the Bristol language development scales was analysed in a fully blinded manner.

Outcome measures

The five primary outcomes were auditory comprehension and expressive language scores, phonology error rate, the score for the Bristol language development scales, and a binary variable indicating whether the child, by 12 months, had improved sufficiently on the single clinical measure on which he or she had entered the study to no longer satisfy that particular criterion. The 11 secondary outcomes were four therapy outcome measures¹⁷; attention and play ratings; the Vineland socialisation domain; three separate binary variables indicating improvement or no deterioration over 12 months in auditory comprehension, expressive language, and phonology; and a fourth binary variable reflecting whether, at the 12 month follow up, the child satisfied any of the clinical criteria (a reassessment of eligibility for the trial irrespective of the initial clinical criterion on which the child entered the trial).

Sample size considerations

For the binary primary outcome of clinical progress it was expected that about 15% of children would improve in the watchful waiting group by the 12 month reassessment. A total of 146 to 166 children (73 to 83 in each arm) was needed to detect a 20% difference between the two arms (that is, 15% v 35%) at a two sided 5% significance level, for 80% to 85% power. This target range of sample sizes had 80% to 85% power to detect differences between the trial groups of 0.43 to 0.50 standard deviations for the continuous outcome measures. Initial assessments continued until the number of children randomised was as close as possible to the upper end of the sample size range.

Analysis

All analyses were performed with the SPSS statistical package (version 10.0), and a two sided 5% significance level was used throughout. The trial arms were compared on an "intention to treat" basis. The continuous outcome measures were analysed by using repeated measures analysis of covariance for the assessments at 6 and 12 months, with adjustment for the baseline assessment of the outcome measurement. The scores for the Bristol language development scales and the Vineland domains were analysed similarly, but with simple analy-



Progress of children through trial

sis of covariance for the assessment at 12 months only. The binary outcome measures reflecting improvements over 12 months from baseline were analysed by using χ^2 tests and logistic regression. In addition, for the two primary outcomes that were not age standardised (the Bristol language development scales and the phonology error rate), the relevant regression models were repeated after adjustment for age.

Results

Participant flow and follow up

In all, 507 children under $3\frac{1}{2}$ years of age were referred to the participating speech and language therapy clinics (figure). Parental consent was obtained for all but 70 of these children. Limited information was available for these 70 children, although their average age (34.5 (range 23 to 42) months) was closely similar to that of the randomised children. The trial assessments were not administered to these children, but, given that 46 of them received therapy outside the trial, over half of this group could potentially have been eligible for the trial.

Of the remaining 437 children, 278 (64%) were not eligible according to the general or clinical criteria: 5 were not new referrals; 9 failed to attend the baseline assessment appointments; 61 did not satisfy the severity criteria (box 2); 134 children were not considered by the therapists to have significant clinical difficulties; 26 had secondary speech and language delays or oromotor, voice, or fluency difficulties; and 43 were excluded owing to various aspects of family background (box 1). The 159 eligible children were subsequently randomised—71 to the therapy group and 88 to the watchful waiting group (figure). Losses to follow up were minimal.

Characteristics at baseline

The children in both trial arms were closely similar for a broad range of baseline characteristics (table 1). Moreover, the mean age of the 159 children randomised was virtually identical to that of the 278 who were ineligible. The ineligible children had better scores on the auditory comprehension and expressive language tests but had a similar mean phonology error rate compared with the randomised children. Compared with population norms for the age standardised scales, at baseline the children in the trial had mean auditory comprehension scores about 1 SD below the mean and expressive language scores nearly 2 SDs below the mean. For the Vineland adaptive behaviour scales, the socialisation score was also about 1 SD below the (normative) mean, whereas the daily living and motor skills scores were both within 1 SD of the mean.

Measures of therapy received

Table 2 shows that the amount of therapy received during the trial by the 71 children allocated to the therapy group was not intensive, with minimal differences between the three clinical criteria. A much higher proportion of the watchful waiting group than of the therapy group (20% v 4%) departed from the protocol, mostly at the 6 month follow up (figure). The 18 children who were allocated to the watchful waiting group but whose parents subsequently requested therapy received an intensity of therapy closely similar to that received by children in the therapy group, but less therapy in total (commensurate with the shorter period of time in which it was delivered).

 Table 1
 Characteristics of children at baseline. Values are numbers (percentage) of children, unless stated otherwise

	Therapy (n=71)	Watchful waiting (n=88)
Maternal education*:		
No qualifications	7 (11)	16 (18)
O level or similar (CSE or technical qualification)	53 (80)	63 (72)
A level and higher	6 (9)	8 (9)
Mean (range) age in months	34.2 (18-42)	34.2 (24-42)
Male	55 (77)	65 (74)
Receiving child care	42 (59)	47 (53)
Diagnosed hearing loss	7 (10)	6 (7)
Mean (range) auditory comprehension score	82.1 (53-118)	83.0 (55-127)
Mean (range) expressive language score	77.3 (59-135)	76.5 (53-104)
Mean phonology error rate (%)	58.1	60.7
Mean (range) score on Bristol language development scale	2.3 (0-8)	2.4 (0-10)
Mean (range) score on Vineland adaptive behaviour scale:		
Daily living skills	89.5 (66-125)	91.9 (67-115)
Motor skills	91.5 (65-115)	91.2 (63-115)
Socialisation skills	82.1 (68-101)	81.8 (79-114)

*Data for some children in the therapy group were missing

Table 2 Measures of therapy received during trial (therapy group, n=71)

	Mean (range)
No of hours of therapy	6.2 (0-15)
No of contacts with therapists	8.1 (0-17)
requency of therapy	Once a month (once a week to once every two and a half months)
No of months over which therapy took place	8.4 (0.9-12)
ength of sessions (minutes)	47 (20-75)

Table 3 Primary outcome measures*

	Difference or odds ratio (95% confidence interval)	P value
Auditory comprehension†	4.1 (0.5 to 7.6)	0.025
Expressive language†	1.4 (-2.1 to 4.8)	0.44
Phonology error rate†	-4.4 (-12.0 to 3.3)	0.26
Bristol language development scale‡	0.1 (-0.4 to 0.6)	0.73
Improvement by 12 months on the clinical criterion on which child entered study§	1.3 (0.67 to 2.4)	0.46

*Data were missing for all measures in both groups: analyses were based on 64 (therapy group) and 80 children (watchful waiting group) for auditory comprehension; 63 and 77 for expressive language; 57 and 62 for the phonology error rate; and 71 and 84 for improvement by 12 months.

†Difference in means obtained from repeated measures analysis of covariance with adjustment for baseline assessment of the outcome measure, with the difference for the therapy group minus that for watchful waiting averaged across the two follow up times.

‡Difference in means obtained from analysis of covariance for 12 month scores adjusted for baseline assessment of the outcome measure.

§Odds ratios obtained from logistic regression.

Table 4 Secondary outcome measures

	Difference or odds ratio (95% confidence	
	interval)	P value
Therapy outcome measures*:		
Impairment	0.1 (-0.2 to 0.5)	0.44
Disability	0.1 (-0.2 to 0.4)	0.56
Handicap	-0.1 (-0.4 to 0.2)	0.67
Wellbeing	0.04 (-0.2 to 0.3)	0.75
Attention level*	0.02 (-0.3 to 0.3)	0.91
Play level*	0.04 (-0.2 to 0.2)	0.70
Vineland socialisation skills†	0.6 (-3.1 to 4.2)	0.76
Improvement by 12 months on auditory comprehension‡	1.4 (0.71 to 2.6)	0.35
Improvement by 12 months on expressive language‡	1.1 (0.61 to 2.2)	0.68
Improvement by 12 months on phonology‡	2.7 (1.2 to 6.3)	0.015
Reassessment of eligibility at 12 months‡	2.1 (1.0 to 4.2)	0.036

*Difference in means obtained from repeated measures analysis of covariance with adjustment for baseline assessment of the outcome measure, with the difference for the therapy group minus that for watchful waiting averaged across the two follow up times.

†Difference in means obtained from analysis of covariance for 12 month scores adjusted for baseline assessment of the outcome measure.

‡Odds ratios obtained from logistic regression.

Table 5 Characteristics of children at 12 month follow up

	Therapy (n=71)*	Watchful waiting (n=84)*
Mean (SD) auditory comprehension score	87.3 (15.9)	84.3 (15.5)
Mean (SD) expressive language score	83.8 (15.1)	81.2 (15.8)
Mean phonology error rate (%)	27.2%	34.4%
Mean (SD) score on Bristol language development scale	4.9 (1.7)	4.8 (1.8)
*Data and a facility of a structure in halfs and a structure in the structure in the structure is the struct		

*Data were missing for all measures in both groups: analyses were based on 71 (therapy group) and 84 children (watchful waiting group) for auditory comprehension; 70 and 82 for expressive language; 70 and 81 for the phonology error rate; and 71 and 84 for the Bristol language development scale.

Therapy provided in the study tended to focus on several areas of language simultaneously. Therapy techniques included Derbyshire language scheme tasks, as well as everyday play and games used as contexts for modelling language for the child. Goals covered a wide range of language stages—for example, understanding and building single words, using narratives, and identifying consonants in words.

Comparisons of trial groups at follow up

For the outcomes available at both 6 and 12 months, the repeated measures analyses of covariance did not show that the differences between the groups changed across the two follow up points (for example, P values for group by time interactions were 0.82, 0.63, and 0.87 for the three relevant primary outcomes in table 3). For all such outcomes the results presented are for the measurements averaged across the two follow up

points. As adjustment for age, where relevant, had no effect, the unadjusted results are presented. For most of the primary variables few data were missing for the analyses in table 3 (most missing values were for the baseline assessment of the outcome measure). Sensitivity analyses that were performed without taking account of adjustment for these baseline assessments did not materially change the findings.

Although all of the observed comparisons for the primary outcome measures were in favour of the therapy group, only one was statistically significant —namely, auditory comprehension (table 3). For this outcome, the difference of 4.1 points corresponds to about 0.3 SD, with the upper confidence limit being about 0.5 SD. For the other three continuous primary outcomes, the confidence intervals allow us to rule out clinically significant differences in favour of the watchful waiting group (that is, differences of >0.25 SD). Conversely, the confidence interval for the binary primary outcome includes important differences in either direction.

For the 11 secondary outcomes none of the seven continuous variables was significant and all of the observed differences were very close to zero, with narrow confidence intervals in most cases (table 4). Of the four binary outcomes, however, two were significant, with a greater proportion of children in the therapy group improving their phonology and no longer satisfying the original eligibility criteria for the trial. Of the 71 children in the therapy group, 27 (38%) were no longer eligible by the end of the trial, compared with 19 (23%) of the 84 children followed up in the watchful waiting group. Overall, 109 (70%) children still satisfied the eligibility criteria at the 12 month follow up.

Descriptive statistics of primary outcomes at 12 months

Table 5 gives descriptive statistics for the primary outcomes at the 12 month follow up. Although the children's mean scores for the standard assessments of auditory comprehension and expressive language improved by 1-6 points over the 12 months, the large standard deviations show that considerable variations remain. In addition, overall the children's phonology error rates improved by about 30 percentage points from an initial 60%.

Discussion

This trial is by far the largest to date investigating the effectiveness of speech and language therapy in preschool children. Improvement in the therapy group was significant (compared with the watchful waiting group) for only one of the five primary outcomesauditory comprehension. Moreover, the two secondary outcomes for which the results were significant measure different aspects. The two possible explanations are, firstly, that the statistically significant findings may simply be due to chance, and, secondly, that there may be a therapeutic benefit across a range of measures, with differential sensitivity resulting in only a small number of (different) outcomes yielding statistical significance. Table 3 supports this latter interpretation, given the direction of the estimates for the primary outcomes and that their confidence intervals generally include large effects in favour of the therapy

group but rule out clinically significant differences in favour of the watchful waiting group. The sizeable minority of parents in the watchful waiting group who requested therapy at the 6 month follow up shows that some parents found it difficult to accept a 12 month period of monitoring.

Limitations of trial

Overall, the impact of therapy in this trial was small, perhaps because of the relatively low level of therapy provided-considerably lower than levels reported in previous studies.³¹⁹ On the other hand, this trial aimed to evaluate routine therapy rather than a prescribed regimen.

In line with the pragmatic design, the children included in the trial presented with a wide range of types and severity of speech and language difficulties. Although the children were stratified according to their broad entry criteria, which ensured similar groups in this respect, the sample size of the clinical groupings was too small to detect significant differential effects.

Blinding was maintained for all baseline assessments and for the language sample at follow up. Although every effort was made to retain blinding at the follow up assessments, in the presence of parents strict blinding was inevitably not always feasible for the other outcomes. The consistency of findings for the various outcomes, however, suggests that this did not seriously bias the results.

Relation to literature

Although the level of therapy in this trial was lower than in smaller scale, more explanatory trials, the present study was a relatively large and randomised trial and it had considerably longer follow up (12 months) than other studies.²⁰⁻²²

Conclusions

Most children in this study still had important clinical difficulties at 12 months, regardless of trial allocation; indeed, many remained eligible for the trial, with little evidence of "spontaneous resolution." This study provides little evidence for the effectiveness of speech and language therapy when compared with "watchful waiting" over 12 months. In clinical terms, these findings suggest that speech and language therapy for preschool children should be reconsidered in terms of appropriateness, timing, nature, and intensity. Further research into more specific types of provision with subgroups of children is required to identify better treatment methods.

We thank the speech and language therapists from Frenchay and United Bristol Healthcare Trusts for their hard work and the families and their health visitors for support and cooperation. The department of social medicine at the University of Bristol is the lead centre of the MRC Health Services Research Collaboration.

Contributors: The original trial design was formulated by PE and TJP, with all authors contributing to its subsequent development. The trial was implemented by SR and MG, with support and advice from TJP and PE. Data management and analysis were carried out by MG, SR, and TJP, under the overall supervision of TJP. MG produced the first draft of the paper, with substantial redrafting by TJP and SR and additional input and academic support from PE. All authors will act as guarantors for the paper.

Funding: Research and development directorate of the South and West regional office of the NHS Executive.

Competing interests: None declared.

What is already known on this topic

A systematic review has shown short term efficacy of speech and language therapy for young children in experimental environments

Evidence is lacking on the long term effectiveness of early intervention for preschool children as provided in a service setting

What this study adds

This study provides little evidence for the effectiveness of speech and language therapy compared with "watchful waiting" over 12 months

Providers of speech and language therapy services should reconsider the therapy offered to preschool children

The low rate of resolution of difficulties suggests that further research is needed to identify effective ways of helping these children

- Law J. The early identification of language impairment in children. London: Chapman and Hall, 1992.
- 2 Hall DMB, ed. Health for all children. Report of the third joint working party on child health surveillance. Oxford: Oxford University Press, 1996.
- 3 Law J, Boyle J, Harris F, Harkness A, Nye C. Screening for speech and language delay: a systematic review of the literature. Health Technology Assessment 1998;2(9):1-184.
- Paul R, Looney SS, Dahm PS. Communication and socialization skills at ages 2 and 3 in "late-talking" young children. J Speech Hear Res 1991;34:858-65.
- Rescorla L, Schwartz E. Outcome of toddlers with specific expressive language delay. *Appl Psycholinguist* 1990;11:393-407. Felsenfeld S, Broen PA, McGue M. A 28-year follow-up of adults with a $\mathbf{5}$
- 6 history of moderate phonological disorder: linguistic and personality results. J Speech Hear Res 1992;35:1114-25. Stothard SE, Snowling MJ, Bishop DVM, Chipchase BB, Kaplan CA.
 - Language-impaired preschoolers: a follow-up into adolescence. J Speech Hear Res 1998;41:407-18.
- Johnson CJ, Beitchman JH, Young A, Escobar M, Atkinson L, Wilson B, language impairments: speech/language stability and outcomes. J Speech Language Impairments: speech/language stability and outcomes. J Speech Lang Hear Res 1999;42:744-60.
- Glogowska M, Campbell R. Investigating parental views of involvement in pre-school speech and language therapy. Int J Lang Commun Disord (in oress).
- 10 Glogowska M. The effectiveness and acceptability of community speech and language therapy provision for pre-school children [PhD thesis]. Bristol: University of Bristol, 1999.
- Zimmerman IL, Steiner VG, Pond RE. Preschool language scale-3. San Antonio: Psychological Corporation, 1992.
 Pagel Paden E, Novak MA, Beiter AL. Predictors of phonologic
- inadequacy in young children prone to otitis media. J Speech Hear Disord 1987;52:232-42.
- 13 Gutfreund M, Harrison M, Wells G. Bristol language development scales. Windsor: NFER-Nelson, 1989. 14 Sparrow SS, Balla DA, Cicchetti DV. Vineland adaptive behavior scales.
- Circle Pines, Minnesota: American Guidance Service, 1984. 15 Cooper J, Moodley M, Reynell J. Helping language development. London:
- Cooper J, Moodue's M, Reynea J. Propage anguage and the second sec
- therapy. San Diego: Singular Publishing, 1997. 18 Roulstone S, Glogowska M, Enderby P, Peters TJ. Issues to consider in the
- evaluation of speech and language therapy for pre-school children. *Child Care Health Dev* 1999;25:141-55.
- 19 Robertson SB, Weismer SE. Effects of treatment on linguistic and social skills in toddlers with delayed language development. J Speech Lang Hear Res 1999;42:1234-47
- 20 Fey ME, Cleave PL, Long SH, Hughes DL. Two approaches to the facilitation of grammar in children with language impairment: an experimental evaluation. J Speech Hear Res 1993;36:141-57.
- 21 Gibbard D. Parent-based intervention with pre-school language delay children. Eur J Disord Commun 1994;29:131-52.
- 22 Almost D, Rosenbaum P. Effectiveness of speech intervention for phonological disorders: a randomised controlled trial. Dev Med Child Neurol 1998;40:319-25.

⁽Accepted 14 June 2000)