Appendix 6 (as supplied by the authors): GRADE evidence profile: effect of treatment on developmental delay outcomes — RCTs*

Quality assessment							No of patients		Effect	Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intervention	Control	SMD / MD (95% CI)		
Effect on language impairment (measured with: objectively; Better indicated by higher values)											
3[1]	randomised trials	serious[2]	no serious inconsistency[3]	no serious indirectness[4]	no serious imprecision[5]	none[6]	117	122	SMD 0.81 higher (0.01 to 1.60 higher)	⊕⊕⊕O MODERATE	CRITICAL
Effect on Adaptive functioning (measured with: objectively; Better indicated by higher values)											
1[7]	randomised trials	serious[8]	no serious inconsistency[9]	no serious indirectness[10]	serious[11]	none[6]	71	84	MD 0.60 higher (-3.05 lower to 4.25 higher)	⊕⊕OO LOW	CRITICAL
Effect on Gross & Fine Motor Skills (measured with: objectively; Better indicated by higher values)											
0[12]	N/A[13]	NA[14]	N/A[15]	N/A[16]	N/A[17]	N/A[18]	-	-	-[19]	N/A[20]	IMPORTANT
Effect on Performance and Cognition											
0[12]	N/A[13]	N/A[14]	N/A[15]	N/A[16]	N/A[17]	N/A[18]	-	-	-[19]	N/A[20]	CRITICAL

GRADE Working Group grades of evidence

High quality: Further research is very unlikely to change our confidence in the estimate of effect.

Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: We are very uncertain about the estimate.

***Reference:** Warren R, Kenny, M, Fitzpatrick-Lewis D, et al. Screening and treatment for developmental delay in early childhood (ages 1-4 years): a systematic review. Calgary: Canadian Task Force on Preventive Health Care; 2016. Available: canadiantaskforce.ca/ctfphcguidelines/2015-developmental-delay/systematic-review/ (accessed 2016 Mar. 29).

[1] Hund-Reid CSP. Effectiveness of phonological awareness intervention for kindergarten children with language impairment. *Can J Speech Lang Pathol Audiol* 2013;37:6-25; Buschmann A, Jooss B, Rupp A, et al. Parent based language intervention for 2-year-old children with specific expressive language delay: a randomised controlled trial. *Arch Dis Child* 2009;94:110-6.
Glogowska M, Roulstone S, Enderby P, et al. Randomised controlled trial of community based speech and language therapy in preschool children. *BMJ* 2000;321:923-6.

[2] Using Cochrane's Risk of Bias tool, for this outcome two studies were rated as unclear risk of bias, one study was rated as high risk of bias. Across studies, there was a lack of certainty (unclear ratings) regarding sequence generation (100%), and allocation concealment (33%) ; and high risk of bias associated with blinding of participants and personnel (100%); incomplete outcome reporting (33%) and other risk of bias (33%; i.e. baseline characteristics, pre-hoc power analysis, sample size >30 per arm). Given that most of the information is from studies at moderate risk, this body of evidence was downgraded for serious study limitations.

[3] The statistical heterogeneity is high [Chi2=12.60, df = 2 (P=0.002); I2=84%] but the direction of the effect is consistent and the confidence intervals overlap across most studies. The statistical heterogeneity is most likely due to small versus large treatment effects observed across studies. This body of evidence was not downgraded for inconsistency.

[4] Three RCTs provided data for this outcome. All studies included mixed gender children with ages ranging from 24.7 months to 66.68 months. In all studies the intervention group received a language intervention delivered in a primary care setting (two studies) or a school setting (one study) by either language therapists, a pediatric neurologist or education assistants. The control group received no intervention. The outcome of language

impairment was accessed using DIBEL, SETK-2 and BLDS cross the three studies. Intervention lengths ranged from 14 weeks to 9 months; follow-up was immediate post in all studies. The studies were conducted in Germany, the UK and Canada. There were no serious concerns regarding indirectness for this body of evidence and it was not downgraded.

[5] The sample size is not adequate i.e. < 300 (117 intervention arm, 122 control arm) but the pooled effect estimate is precise and confidence interval does not include the null value "0" [SMD= 0.81 (0.02, 1.60)]. This body of evidence was not downgraded for serious concerns regarding imprecision.

[6] There were too few studies (n<10) to assess publication bias.

[7] 1) Glogowska 2000

[8] Using Cochrane's Risk of Bias tool, for this outcome the study was rated as unclear risk of bias. In this study, there was a lack of certainty (unclear rating) regarding sequence generation; and high risk of bias associated with blinding of participants and personnel. Given that most of the information is from a study a moderate risk, this body of evidence was downgraded for serious study limitations

[9] The statistical heterogeneity across studies could not be assessed due to only one study providing data for this outcome.

[10] One study provided data for this outcome. The study included mixed gender children with a mean age of 34.2 months. The intervention consisted of one-on-one speech and language therapy with trained speech and language therapists, over 8.4 months. The outcome of Social and personal activities of daily living (Socialization, adaptive functioning) was accessed using VABS. Follow-up was immediate post. The study was conducted in the UK and was published in 2000. There were no serious concerns regarding indirectness for this body of evidence and it was not downgraded.

[11] The sample size is not adequate i.e. < 300 (71 intervention arm, 84 control arm) and the pooled effect estimate is not precise with confidence interval including the null value "0" [MD= 0.60 (-3.05, 4.25)]. This body of evidence was downgraded for serious concerns regarding imprecision.

[12] No studies were found that met the inclusion criteria of this review for this treatment or outcomes

[13] No studies were found that met the inclusion criteria of this review for this treatment or outcomes

[14] Risk of bias cannot be assessed

[15] Inconsistency cannot not be assessed

[16] Inconsistency cannot not be assessed

[17] Inconsistency cannot not be assessed

[18] Inconsistency cannot not be assessed

[19] Inconsistency cannot not be assessed

[20] Since there no studies the overall quality of the evidence cannot be determined