A global perspective of the prevalence of low language in children: a scoping review and evidence and gap map protocol

Lisa Irene Jones¹
Michelle A Krahe²
Amy Giesberts¹
Martin Downes¹
Robert S Ware¹
Sheena Reilly¹

1. Griffith Health, Griffith University, Gold Coast, Queensland Australia
2. College of Medicine and Dentistry, James Cook University, Smithfield, Queensland Australia

Corresponding author:

Lisa Irene Jones
lisa.jones@griffith.edu.au

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Objective: The objective of this scoping review is to identify, categorize, and map the existing evidence of low language prevalence in children aged 4 to 18 years globally. Additionally, it seeks to explore associated factors, and map evidence gaps to improve future reporting practices.

Introduction: The study of language difficulties in children has evolved over the years, marked by changing terminologies and criteria. Understanding why prevalence varies is critical to address the need for accurate prevalence estimates, highlight factors influencing reported frequency, and inform policy, resource allocation, and research priorities.

Inclusion criteria: This review will consider studies that report low language prevalence for children aged 4 to 18 years from any country that are derived from studies or epidemiological samples in community/population settings, including households, birth registers/cohorts, and schools. To encompass the changing use of terminologies and criteria, the term low language is used as an umbrella term to describe children whose language development is behind their peers.

Methods: This review protocol is guided by the Joanna Briggs Institute methodology for scoping reviews. Key databases will be systematically searched for English-language studies, published between 01 January 1970 to 30 June 2023. Two independent reviewers will screen titles and abstracts, while three reviewers will assess full texts. The final selection of studies will undergo review by a broader expert group and data extraction will be conducted by one reviewer and verified for accuracy. The findings will be presented in narrative format, including tables and figures to aid in data presentation.

Review registration: Open Science Framework (https://doi.org/10.17605/OSF.IO/MKNS2)

Keywords: language; children; prevalence; epidemiological samples; scoping review
Introduction

Whilst not all children learn language at the same rate, it is well recognized that a group of children have problems acquiring language. The first descriptions of children having such difficulties appeared nearly two centuries ago(1), and since that time there has been a proliferation of terms for language difficulties, with some used briefly, whereas others have gained greater traction(2). A pivotal debate in 2014 prompted an international consensus panel to examine the terminology and criteria used to describe language problems in children(3, 4). This resulted in the endorsement of the term Language Disorder (LD) to reflect “... a profile of difficulties that causes functional impairment in everyday life and is associated with poor prognosis”(5, p1068) and the term Developmental Language Disorder (DLD) to describe children who have difficulties learning language in the absence of any known biomedical etiology(5, 6).

The lack of agreement about the terminology and criteria used to identify children with language problems has been described as detrimental to research(5) including our understanding of the prevalence of the condition. Estimates remain imprecise, ranging from 7-17% between cohort studies(7-9). In a recent systematic review, Hill, Calder (10) reported even wider estimates, from 0.4% to 25.2% in children aged 1 to 18 years, with greater stability of language level observed in children aged 5 years and older. We note that these findings are not specific to the study of childhood language; similar variation in prevalence has been observed in other developmental conditions including autism, attention-deficit/hyperactivity disorder (ADHD). Further, reports demonstrated that some of the factors that influenced prevalence estimates, included the child age at which prevalence was estimated, the measure/s used, the definition and diagnostic criteria of the disorder, and the geographic location and population of the sample(11).

Accurate prevalence estimates are essential for informing public policy, raising awareness, and developing research priorities(12). They are instrumental in resource allocation, influencing funding decisions, and the expansion of service provision. Therefore, understanding why current prevalence estimates vary is important. There are likely to be multiple factors that influence estimates including the study design, the definition and measurement of language problems, and the age of the child at assessment. As in the study of other developmental conditions, the geographical location may also play a role. Identifying these factors could influence precision in current reporting practices and inform future research directions. A global estimation of prevalence serves as a valuable tool for advocacy and guides research focus, particularly in regions where prevalence data is limited or unavailable(13).
At the outset we recognized that this review would need to consider the changing terminology and criteria used to describe children with language problems in the search strategy. After Hill, Calder (10) we adopted the term ‘low language’ to ensure we captured a wide range of reports on language ability in children. The authors limited the terms they used to describe the number of children with language problems to prevalence, incidence, and epidemiology and may have inadvertently missed published studies that either did not use these terms or where reports of low language prevalence were embedded in broader published reports, such as in nationally representative birth cohorts. We therefore adopted broad search criteria to capture published reports from a range of studies that might have commenced at any time over the past 40-50 years, during which terminology and criteria will have changed.

We anticipate encountering numerous studies reporting on the prevalence or incidence of low language worldwide as well as variability in the study contexts, including the type of sampling and its derivation, as well as variation in how low language was determined. This encompasses considerations such as when and how measurements were taken, including the age of the child at measurement, the tools used for measurement, the language domains measured and the reporting methods employed (e.g., total scores, subgroup scores). Furthermore, we anticipate encountering studies reporting on comorbidities associated with low language.

To our knowledge, there appear to be no formal reviews of factors associated with the prevalence of low language in children, globally. The objective of this review is to provide an overview of the evidence related to low language prevalence in children, and to identify, map and categorize factors associated with prevalence reporting. The findings from this review will have several practical implications. Firstly, it will provide a comprehensive understanding of the global prevalence of low language in children aged 4 to 18 years, thereby informing policymakers, healthcare professionals, and educators about the scope of this issue. Secondly, by exploring factors driving the prevalence within the studies, such as population age, measurement methods, and geographic location, the study can identify gaps in the current literature and inform low language prevalence reporting in future studies. Finally, mapping evidence gaps will help identify areas where further research is needed, facilitating the improvement of reporting practices and ultimately leading to more accurate prevalence estimates and better-informed decision-making in healthcare and education sectors.
Review question(s)

1. What are the key characteristics of studies that report prevalence of low language in children aged 4 to 18 years?
2. What are the factors associated with the reporting of prevalence?
3. What are the gaps in the current literature, and how do they, and the findings of this scoping review, inform low language prevalence reporting in future studies?

Inclusion criteria

Participants

This review will consider studies that include children aged 4 to 18 years old. Four years was selected as the lowest age limit because language development in the early years is characterized by greater fluctuation and stabilizes by 4 years of age.

Concept

This review will consider studies that describe or evaluate the prevalence of low language, which is used to describe a group of children whose language development is behind their peers. This can include terms such as developmental language disorder, specific language impairment or language delay. This concept search was designed using key terminology for language disorders and guided by expert researchers within the field. Low language is determined by a specific language measure(s) such as standardized language assessment tools, observational methods (i.e., language samples collected during naturalistic interactions), parent or teacher reported measures, or a combination of multiple measures that provide a comprehensive evaluation of language abilities. Language outcome determined by report only, clinical referral, or ascertained from a clinical record will be out of scope.

Context

This review will consider studies where data is derived from studies or epidemiological samples in community or population settings, including households, birth registers/cohorts, and schools. This includes cross-sectional, prevalence, incidence, cohort, demographic, proportion, epidemiology, or population studies. Excluded will be studies involving clinical, referred or volunteer-based samples, and samples assembled retrospectively. Studies that
lack sufficient information to estimate prevalence or those devoid of original data, such as literature reviews or commentaries, will also be excluded from consideration.

**Types of sources**

This review will consider studies that have been conducted in any geographic location or setting. Both experimental and observational study designs, including randomized controlled trials, non-randomized controlled trials, before and after studies, and cohort or population studies will be considered.

**Methods**

The proposed scoping review will be conducted in accordance with the Joanna Briggs Institute (JBI) methodology for scoping reviews (17, 18). The review will report findings following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) (19). This protocol has been registered with the Open Science Framework ([https://osf.io/mkns2](https://osf.io/mkns2)).

**Search strategy**

The search strategy will aim to locate both published primary studies, and reports of primary studies. An initial limited search of MEDLINE (via Ovid) was undertaken to identify articles on the topic. The text words contained in the titles and abstracts of relevant articles, and the index terms used to describe the articles, were used to develop a full search strategy for MEDLINE (via Ovid) in Appendix I. Collaboration with a member of the Griffith University Library ensured the accuracy and comprehensiveness of the search strategy, which will be adapted for each included information source. The final search strategy will be run on various databases, including MEDLINE via Ovid, ERIC via ProQuest, Psycinfo via Ovid, Web of Science, Scopus and CINAHL via EbscoHost and the reference lists of articles selected for inclusion will be screened for additional papers. Only articles published in English between 01 January 1970 to 30 June 2023 will be included in the final review.

**Study selection**

Following the search, all identified records will be collated and uploaded into the review management software Covidence (Veritas Health Innovation, Melbourne, Australia). The software will automatically remove duplicates, and this will be checked to ensure accuracy. Following an initial pilot test, two independent reviewers will screen titles and abstracts to
assess their alignment with the inclusion criteria. Potentially relevant papers will undergo full retrieval, imported into Covidence, and assessed in detail against the inclusion criteria by three independent reviewers. Any full-text papers not meeting the criteria will be recorded and reasons will be included in the scoping review. Disagreements among reviewers will be resolved through discussion. The final selection of studies will undergo review by a broader expert group. The search results will be presented in the final scoping review and depicted in a flow diagram following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) guidelines (19, 20).

Data extraction

Data will be extracted by one reviewer using a data extraction tool developed by the reviewers. The data extraction tool will be tested independently by three reviewers on three studies to ensure that specific details about the population, concept, context, and study methods of significance, to ensure consistency in extraction. The draft data extraction tool is provided in Appendix II and will be modified and revised as necessary during the extracting process. Modifications will be detailed in the full scoping review. Any disagreements that arise between the reviewers will be resolved through discussion. Authors of papers will be contacted to request missing or additional data, where required.

Data analysis and presentation

Studies included in this scoping review will be analyzed using quantitative descriptive analysis and will be visualized through evidence and gap maps. The characteristics of studies reporting the prevalence of low language in children, the types of samples involved (i.e., countries, cohorts, language disorder), along with their respective study design and key language outcomes will be described. Additionally, this review will provide an analysis of factors associated with prevalence, categorizing them as evidence for positive, negative, or neutral influence. Potential variations in the definitions of low language across studies will undergo quantitative content analysis, utilizing inductive categories. A bibliographical analysis will be used to conduct content analysis of the current research evidence, capturing prominent research themes in the field, and identifying evidence gaps. Knowledge translation activities will include publication in a peer-reviewed journal and conference presentations. Through direct collaboration with global, national and local networks including research collaborations and partnerships, professional associations, clinical groups, organizations responsible for policy setting and resource allocation and through organizations for people with language disorder and their families.
Acknowledgements

We would like to thank the Early Language in Victoria Study (ELVS) team of investigators, past and present, who have provided input.

Funding

This study is supported through National Health and Medical Research Council Clinical Trials and Cohort Studies (NHMRC CTCS) APP1184250; Centre Research Excellence (CRE) APP2015727.

Author contributions

LIJ, MAK, AG & SR contributed to the initial conceptualization, drafting and finalization of this manuscript. MD and RSW made substantial contributions to the finalization of the protocol. All authors have reviewed and approved the complete manuscript.

Conflict of interest

The authors declare no conflict of interest.
## Appendix I: Search strategy

### MEDLINE (Ovid)

Search conducted on 3 July 2023

<table>
<thead>
<tr>
<th>PCC framework</th>
<th>#</th>
<th>Keywords and Medical Subject Headings (MESH) terms</th>
<th>Records retrieved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population: Children aged 4-18 years old</td>
<td>1</td>
<td>(exp Young Adult/ OR exp Birth Cohort/ OR exp Child/ OR exp Child, Preschool/ OR exp Adolescent/)</td>
<td>3,816,086</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>limit 1 to yr=&quot;1970 -Current&quot;</td>
<td>3,605,150</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>(&quot;young adult&quot; or &quot;birth cohort&quot; or child* or preschool* or pre-school* or kindergarten* or youth or adolescent* or teen*).ti,ab.</td>
<td>1,723,065</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>limit 3 to yr=&quot;1970 -Current&quot;</td>
<td>1,641,013</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>2 OR 4</td>
<td>3,944,424</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Language Development Disorders/ or Specific Language Disorder/</td>
<td>7,331</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>limit 6 to yr=&quot;1970 -Current&quot;</td>
<td>7,297</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>(&quot;language development disorder&quot; or &quot;specific language disorder&quot; or &quot;specific language impairment&quot; or &quot;developmental language disorder&quot; or &quot;expressive receptive language disorder&quot; or &quot;primary language impairment&quot; or &quot;language learning impairment&quot; or &quot;language delay&quot; or &quot;language disorder&quot; or &quot;language impairment&quot; or &quot;speech and language impairment&quot; or &quot;low language&quot; or &quot;primary language impairment&quot;).ti,ab.</td>
<td>8,273</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>limit 8 to yr=&quot;1970 -Current&quot;</td>
<td>8,209</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>7 OR 9</td>
<td>12,464</td>
</tr>
<tr>
<td>Context: Type of study</td>
<td>11</td>
<td>exp Cross-Sectional Studies/ or exp Prevalence/ or exp Incidence/ or exp Cohort Studies/</td>
<td>3,223,996</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>limit 11 to yr=&quot;1970 -Current&quot;</td>
<td>3,203,696</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>(&quot;cross sectional stud&quot; or &quot;cross-sectional stud&quot; or prevalence or &quot;prevalence stud&quot; or incidence or &quot;incidence stud&quot; or &quot;cohort stud&quot; or &quot;demographic</td>
<td>4,862,091</td>
</tr>
<tr>
<td>Step</td>
<td>Query</td>
<td>Results</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>limit 13 to yr=&quot;1970 -Current&quot;</td>
<td>4,823,384</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>12 OR 14</td>
<td>6,379,905</td>
<td></td>
</tr>
<tr>
<td>Final Search</td>
<td>5 AND 10 AND 15</td>
<td>3,354</td>
<td></td>
</tr>
</tbody>
</table>
## Appendix II: Draft data extraction instrument

<table>
<thead>
<tr>
<th>Category</th>
<th>Data to be extracted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Article information</td>
<td>• Author(s) and year of publication&lt;br&gt;• Year language data collected&lt;br&gt;• Study location</td>
</tr>
<tr>
<td>Study design</td>
<td>• Aim(s) of the study&lt;br&gt;• Population and sampling (type and how derived)</td>
</tr>
<tr>
<td>Participant characteristics</td>
<td>• Age at recruitment (range, mean, and SD)&lt;br&gt;• Gender (distribution/ratio)&lt;br&gt;• Socioeconomic status&lt;br&gt;• Comorbidity examined</td>
</tr>
<tr>
<td>Study stage and sample</td>
<td>• Sample size&lt;br&gt;• Sample characteristics, including how sample was ascertained (inclusion/exclusion criteria if applicable)</td>
</tr>
<tr>
<td>Assessment details</td>
<td>• Age at assessment&lt;br&gt;• Screening/assessment method (how/which measures administered)&lt;br&gt;• Diagnostic tool (and/or criteria) and how outcome on tool/criteria was reported (i.e., total score, subscale score, subgroups)&lt;br&gt;• Language domain assessed in the measures (i.e., vocabulary [expressive/receptive], pragmatics, narrative, semantics, syntax, morphology, omnibus, composite)</td>
</tr>
<tr>
<td>Low language identification</td>
<td>• Terminology used to identify language disorder (i.e., SLI, DLD, language difficulty)&lt;br&gt;• How disordered language is determined (i.e., school data, clinical assessment, score on measure)&lt;br&gt;• Prevalence of language difficulty as specified</td>
</tr>
</tbody>
</table>
References


