Research in Developmental Disabilities 31 (2010) 943–950

Contents lists available at ScienceDirect

Research in Developmental Disabilities

Review

Problems and limitations in studies on screening for language delay

Mårten Eriksson [a,*], Monica Westerlund [b], Carmela Miniscalco [c]

[a] Department of Social Work and Psychology, Faculty of Health and Occupational Studies, University of Gävle, Gävle, Sweden
[b] Central Unit for Child Health Care, Department of Women's and Children's Health, Uppsala University, Children's Hospital, Uppsala, Sweden
[c] Division of Speech and Language Pathology, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

ARTICLE INFO

Article history:
Received 16 April 2010
Accepted 22 April 2010

Keywords:
Screening
Screening methodology
Language screening
Language delay
Verification bias
Gold standard

ABSTRACT

This study discusses six common methodological limitations in screening for language delay (LD) as illustrated in 11 recent studies. The limitations are (1) whether the studies define a target population, (2) whether the recruitment procedure is unbiased, (3) attrition, (4) verification bias, (5) small sample size and (6) inconsistencies in choice of "gold standard". It is suggested that failures to specify a target population, high attrition (both at screening and in succeeding validation), small sample sizes and verification bias in validations are often caused by a misguided focus on screen positives (SPs). Other limitations are results of conflicting methodological goals. We identified three such conflicts. One consists of a dilemma between unbiased recruitment and attrition, another between the comprehensiveness of the applied gold standard and sample size in validation and the third between the specificity of the gold standard and the risk of not identifying co-morbid conditions.

Contents

Early identification of most medical problems is advantageous in that it generally provides better opportunities for successful intervention. Yet, medical screening involves many considerations of an economic, ethical and methodological nature before early identification within a certain population can be decided. This study reviews some recent studies on screening for language delay (LD) and identifies some common methodological limitations. Hence, the present context is screening for LD but the methodological considerations are of a general nature.

* Corresponding author at: Department of Social Work and Psychology, Faculty of Occupational Studies, University of Gävle, SE-801 76 Gävle, Sweden. Tel.: +46 26 64 82 11.
E-mail address: marten.eriksson@hig.se (M. Eriksson).

0891-4222/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ridd.2010.04.019

Table 1
Common limitations in screening studies of language delay in children 36 months of age or under.

Study	Author	Target group (expr/contr)[a]	Recruitment procedure	Attrition (%) (expr/contr)[a]	Sample size (expr/contr)[a]	Form of screening
I	Buschmann et al. (2008)	Not defined	Routine health check-up	–	153	Parent questionnaire
II	Eriksson, Westerlund, and Berglund (2002)	1155 (1519)	Routine health check-up	12 (–)	1021	Parent questionnaire
III	Klee et al. (1998)	582	Birth announcement	47	306	Parent questionnaire
IV	de Koning et al. (2004)	5734 (4621)	Routine health check-up	45 (–)	3147	Parent interview
V	Laing, Law, Levin, and Logan (2002)	247 (376)	Routine health check-up	7 (6)	229 (353)	Parent interview
VI	McGinty (2000)	Not defined	Routine health check-up	–	200	Language test
VII	Miniscalco Mattson, Mårild, and Pehrsson (2001)	Not defined	Routine health check-up	–	105	Language test and parent interview
VIII	Rescorla and Alley (2001)	Not defined	School lists	–	422	Parent questionnaire
IX	Sachse and Von Suchodoletz (2008)	1490	Birth announcement	37	932	Parent questionnaire
X	Stott, Merricks, Bolton, and Goodyer (2002)	2590	Community register	28	1861	Parent questionnaire
XI	Westerlund and Sundelin (2000a)	2400	Routine health check-up	2	2359	Language test and parent interview

[a] We use the terminology of experimental (expr) and control (contr) groups, although the "control" may consist of alternative screening.

A major problem in screening for LD is that LD is a matter of degree, whereas the screening method is developed for identification of absolute conditions. Hence, there is not only a problem to decide on a proper "gold standard", but there is also a problem in deciding on cut-off scores. This decision is crucial because it affects the presumed prevalence of LD and hence the required power, sample size, sensitivity, specificity of the screen, etc. Unfortunately, decisions on cut-offs are typically arbitrary and vary widely among studies (see Spaulding, Plante, & Farinella, 2006 for a detailed discussion on this problem).

In a systematic review of the literature on early screening for speech and LD published before 1996, Law, Boyle, Harris, Harkness, and Nye (1998) concluded that although evidence is available for effective interventions for many areas of speech and LD, there is not sufficient evidence of well functioning screening tests and screening procedures to introduce formal screening. The conclusion rests in part on inconsistencies in methodologies of published screening studies that defy clear interpretations. Law et al. (1998) further noted a large variation in the applied validation procedures, i.e. different gold standards were applied.

The present article focuses on six methodological problems and limitations in studies on early language screening for LD published after 1995. The limitations are (1) whether the studies define a target population, (2) whether or not the recruitment procedure is unbiased, (3) attrition, (4) verification bias, (5) small sample size and (6) inconsistencies in choice of gold standard. The selection of studies does not constitute an exhaustive review but each illustrates one or more of the methodological problems described here and how the problems sometimes conflict with each other. All studies concern children 3 years of age or younger at the time of screening. The studies are regarded as "good research" within the area and are published in established journals. Our own studies are included and their strengths and weaknesses discussed. Screening studies without validation are exceptionally limited in that evidence of what they screen for, if anything, is lacking. Therefore, only studies on screening instruments that comprise a validation are included. However, a validation may be concurrent or prospective and the result from the validation may or may not be presented in the same article as the one presenting the initial screen. Table 1 presents the screening studies referred to in Greek numerals and Table 2 presents the validation studies ordered the same way, regardless of whether they are published in the same articles or in subsequent articles. We also refer to the Greek numerals in the text to ease identification.

1. Methodological limitations in screening for LD

We consider a prospective cohort design as most appropriate for evaluations of early language screening. Such designs may include comparisons of more than one screening method (Studies II and V) or comparisons with an unscreened group (Study IV). All the reviewed studies seem to apply some type of cohort design, although some studies are influenced by other designs and the difference is never discussed explicitly. In screen positive designs only the screen positives (SPs) undergo validation. Such designs are only motivated when the gold standard is invasive and hence associated with increased health hazards. That is not the case for gold standards associated with LD. There are also case control studies in which the screen is administered to already diagnosed cases and a group of controls as appropriate for early language screening. Such studies often include an equal number of cases and controls. The major drawback with this design is that it relies on already identified cases, whereas the purpose of a screening programme is to find undiagnosed children at risk of developing LD.

1.1. Target population

To define a *target population* is critical for screening (e.g., Anderman, Blancquaert, Beauchamp, & Déry, 2008). The target population may be either an unselected group formed on a birth register or some such (general screening), or a subgroup