Adams, Valerie; Mathisen, Bernice; Baines, Surinder; Lazarus, Cathy; Callister, Robin. "A systematic review and meta-analysis of measurements of tongue and hand strength and endurance using the Iowa Oral Performance Instrument (IOPI)" Dyshpagia Vol. 28, Issue 3, p. 350-369 (2013)

Available from: http://dx.doi.org/10.1007/s00455-013-9451-3

Chapter 3: A systematic review and meta-analysis of measurements of tongue and hand strength and endurance using the Iowa Oral Performance Instrument (IOPI)

3.1 Introduction

In the early1990s, new tools to measure the pressure generated by contact between the tongue and palate were developed which offered speech-language pathologists an objective means of assessing tongue strength and endurance. One such tool was the IOPI, which has been used primarily in the USA over the past two decades. The IOPI was originally developed to examine the relationships between tongue strength or endurance and speech motor control, and has subsequently been extended to examine relationships with swallowing. Over this time a number of research studies have been conducted using the IOPI on both healthy and clinical populations, which provide data that can be used to establish normative IOPI values for tongue strength and endurance, as well as to investigate the possible influences of age, sex and medical condition on these values [51-53,38,54,8,55-62,49,63-75,37].

The IOPI is a portable, hand-held device that uses an air-filled pliable PVC tongue bulb (approximately 3.5cm long and 1.2cm in diameter (with an approximate internal volume of 2.8ml) connected via an 11.5cm clear PVC tube to measure peak pressure exerted on the tongue bulb measured in kilopascals (kPa). It contains pressure-sensing circuitry, a peak-hold function, and a timer. Researchers have used this device in many studies to measure tongue strength and endurance with excellent inter-rater reliability [75,37]. Currently it is one of the most commonly used measurement techniques available to objectively measure tongue strength and endurance [76]. A hand bulb has also been developed for use with the IOPI, which provides a means of assessing hand in addition to tongue strength and endurance.

3.1.1 Aims and objectives

The primary aim of this systematic review was to evaluate the utility of the IOPI as an effective tool for assessments of both tongue and hand strength and endurance in healthy and clinical populations, and if possible, to identify representative values of these measures. Secondary aims were to investigate the effects of age and sex on the measured values, the impact of clinical conditions, and to determine the use of the IOPI as an intervention tool to improve tongue strength and/or endurance. Meta-analyses to consolidate these effects were conducted where appropriate.

3.2 Methods

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [77] and the Consolidated Standards of Reporting Trials (CONSORT) Statement [78] guided the conduct and reporting of this review.

3.2.1 Eligibility criteria

A systematic computer-based search of 21 databases (Table 3.1) and Google Scholar was conducted for the period between January 1990 and April 2012. The search terms used were: "*Iowa Oral Performance Instrument*" or "*IOPI*". The search was limited to publications in English and peer-reviewed journals. An additional search of the databases using "*tongue strength*" was conducted to ensure maximum inclusion of potential articles. All reference lists in selected journal articles were screened for further potentially relevant articles that met the eligibility criteria. The first authors of two relevant journal articles [69,37] were contacted in April and June 2012 to obtain participant numbers, sex balance, and standard deviations from those studies to allow them to be included in the review. Eligible studies included cross-sectional, time series, prospective cohort, and randomised controlled studies that provided values for tongue or hand strength or endurance measured by the IOPI, or evaluated the IOPI as an intervention tool on measures of strength/endurance in healthy or disordered populations. Exclusion criteria were studies that did not use the IOPI as a measurement device; were abstracts, theses, posters or conference papers; or contained no relevant data.

Table 3.1
A systematic computer-based search of electronic databases and vendors

Cochrane Library (Wiley InterScience)

CINAHL

EBSCO (Academic Search Complete, Communications & Mass Media Complete, Education Resources Complete, Health Source: Nursing, Masterfile Premier, Psyc & Behavioural Sciences Collection, SportsDiscus)

Embase (Elsevier)

Linguistics Language Behavior Abstracts (LLBA)

Medline

OVID

Proquest

PubMed

ScienceDirect

Scopus

Springerlink

Taylor & Francis

Web of Knowledge (Science Citation Index; Social Science Citation Index)

3.2.2 Study selection

After duplicates were deleted, eligibility assessment was performed independently in an unblinded standardised manner by the first author (VA), with any uncertainties resolved by a second author (RC). Retrieved records were screened for relevance and inclusion by title and abstract.

3.2.3 Data extraction process and data items

All data were extracted from the studies by one author (VA). If available, statistics such as 95% confidence interval (CI) or standard error (SE) were converted to the required form (mean ± standard deviations (SD) according to the calculations outlined in the Cochrane Handbook for Systematic Reviews of Interventions (Sections 7.7 and 16.1.3.2) [79]. Information extracted included: (1) authors and year of publication; (2) setting; (3) groups if appropriate; (4) number of participants; (5) sex; (6) mean age; (7) age range; (8) means and standard deviations (SD) of IOPI measures; (9) outcomes of any comparisons between groups and whether *p* values were reported; (10) effect size of any comparisons; and (11) a clear population description (healthy or with disorders).

Studies that were published post-2000 used the second-generation IOPI tongue bulbs (soft vinyl blue silicone bulbs attached to a polyethylene tube, with a 2mm inside diameter). Studies measuring tongue strength published prior to 2000 were further examined to determine bulb texture and colour. Because of slightly different internal volumes and surface areas, pressure values obtained from first-generation clear air-filled tongue bulbs or latex bulbs must be multiplied by 0.87 to be comparable to the present data [70]; this correction was made where required to the values reported in this review. Whether this correction adequately addresses all variations in the materials in the early years is uncertain.

3.2.4 Risk of bias in intervention studies

Risk of bias was assessed for randomised controlled trials and prospective cohort studies by two authors (VA and RC) using a 10-item quality checklist adapted from the Consolidated Standards of Reporting Trials (CONSORT) statement [80]. In the case of disagreement, discussion took place until a consensus was reached. The items and explanations of the scoring for each item are reported in Table 3.2. Each item was scored with a '1' for 'yes' or '0' for 'no'. The studies were then classified as having a low (score ≥ 6) or high risk of bias (score ≤ 5).

3.2.5 Summary measures and synthesis of results

The primary outcome measures for this review were the means ± SD of the IOPI measures (tongue and hand strength [kPa] and endurance [seconds]) for the described population samples. Differences between population groups and the effects of intervention studies were examined using statistical comparisons, and effect sizes such as Cohen's *f*. Meta-analyses of healthy participants with outcomes for tongue strength (kPa) and tongue endurance (seconds) were conducted on eligible evaluation studies. Results were pooled in separate meta-analyses using RevMan 5.1.4 for Windows. All data were continuous and reported on the same scale for age and sex. The aggregate result was calculated as the weighted mean difference (WMD) between age and/or sex. Funnel plots to assess publication bias were generated if greater than 10 studies were included in the meta-analyses [79]. Meta-analysis was deemed inappropriate if results from fewer than three studies were compatible for analysis.

Table 3.2

A 10-item quality checklist scale and explanation of scoring for randomised control trials

Indicator	Quality marker
Study design	Controlled trial * Cohort Study Retrospective case control or single-subject design Case series Case study
Blinding	Assessors blinded * Assessors not blinded or not stated
Sampling/allocation	Random sample adequately described * Random sample inadequately described Convenience sample adequately described Convenience sample inadequately described or hand-picked sample or not stated
Group/participant comparability	Groups/participants at baseline on important factors (between-subject design) or participant(s) adequately described (within-subject design) * Groups/participants not comparable at baseline or comparability not reported or participant(s) not adequately described
Outcomes	At least one primary outcome measure is valid and reliable * Validity unknown, but appears reasonable; measure is reliable Invalid and/or unreliable
Significance	*p* value reported or calculable * *p* value neither reported or calculable
Precision	Effect size and confidence interval reported or calculable * Effect size or confidence interval, but not both, reported or calculable
Intention to Treat (controlled trials only)	Analysed by intention to treat * Not analysed by intention to treat or not stated

*Indicates highest level of quality

3.3 Results

3.3.1 Study selection

A search across 21 databases yielded a total of 295 articles that were identified for inclusion in the review (Figure 3.1). An additional 47 articles were identified from searching the reference lists of included articles. After adjusting for duplications, 162 remained. Of these, 126 studies were excluded, as they did not meet the eligibility criteria. The full texts of the remaining 42 articles were examined in greater detail. Four of these articles did not meet the inclusion criteria, as they did not provide IOPI data on tongue or hand strength or endurance. Thirty-eight studies met the inclusion criteria and were included in the systematic review.

3.3.2 Study characteristics

Of the 38 included studies, 36 were conducted in the United States; one in Brazil and one in Taiwan. The collective sample size was 1729 participants with 882 males (51%) and 847 females. Participants consisted of 53% healthy people and 47% from disordered populations (Parkinson’s Disease (PD), head and neck cancer (HNC), Multiple Sclerosis (MS), Motor Neuron Disease (MND), traumatic brain injury (TBI), nasopharyngeal cancer (NPC), oculopharyngeal muscular dystrophy (OPMD), cerebrovascular accident (CVA), Developmental Apraxia of Speech (DAS), Developmental Verbal Dyspraxia (DVD). The majority of participants were recruited from the community (24%); clinics (21%); no setting stated (21%); hospitals (16%); schools or university (13%); or from other research projects (5%). Age ranges included children and adolescents (3 to 17 years) and adults (18 to 96 years). Included studies were classified as evaluation studies 87% (n = 33) or intervention studies 13% (n = 5).

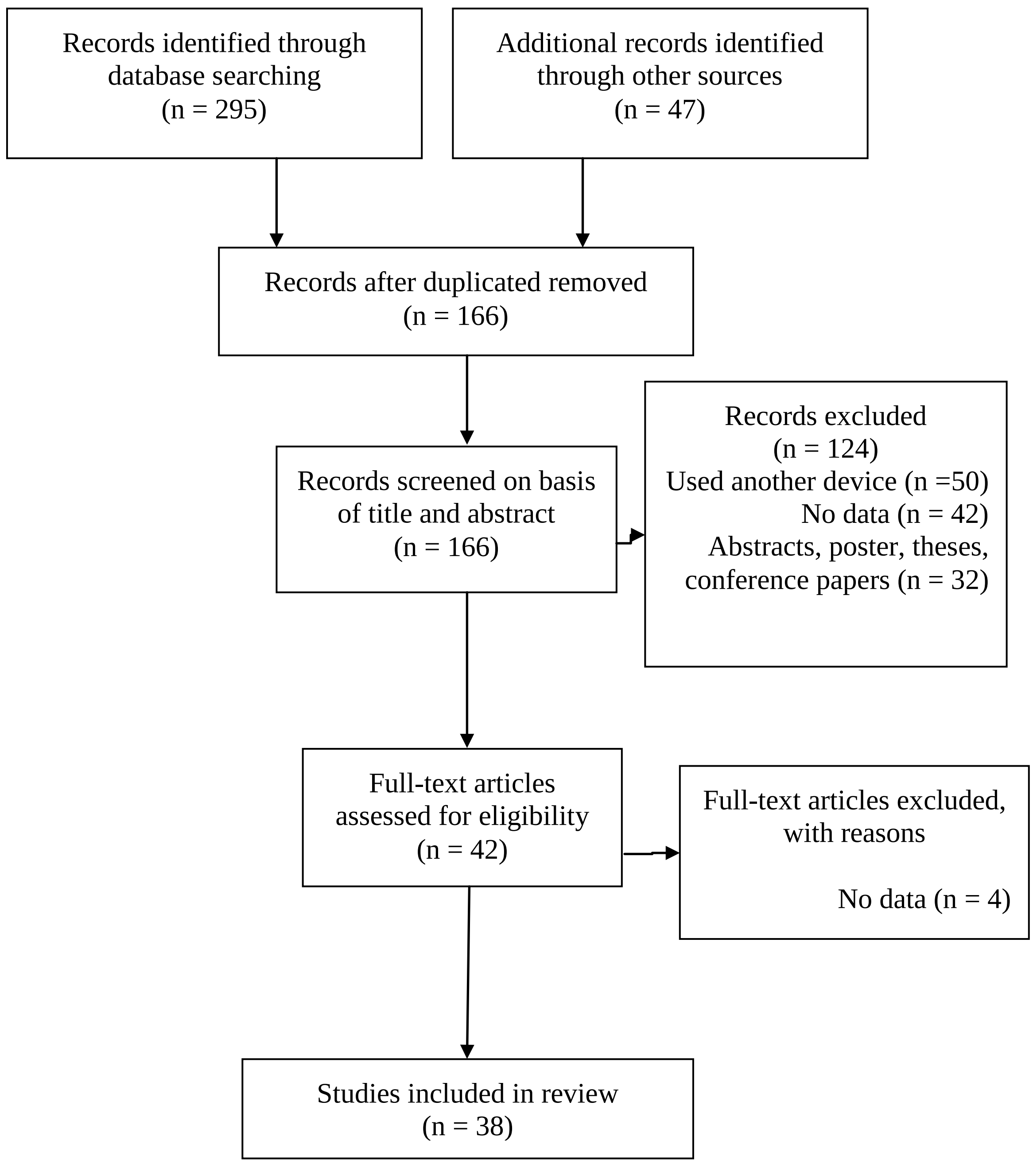

Figure 3.1. A flowchart of the literature search of databases

3.4 Evaluation studies

3.4.1 Tongue strength in healthy populations

Sixteen studies (adults n = 14 and children n = 2) reported measures of tongue strength (in kPa) in healthy individuals (Table 3.3). Mean values ranged from 43 to 78 kPa in healthy adults. Twelve studies reported data for healthy adult males and females; mean values for tongue strength in healthy males ranged from 49.25 ± 18.64 to 73.33 ±12.03 kPa compared to moderately lower values for healthy females (37.00 ±11.36 to 66.96 ±11.60 kPa) at similar ages. Values of tongue strength in the healthy adult population have been reported primarily for anterior elevation and secondarily for posterior elevation. Reports of other tongue strength measures using the IOPI (i.e. lateralisation and protrusion) were not considered for this review. Three studies [8,53,54] measured tongue strength in both the anterior and posterior positions. Two studies [53,8] investigated tongue strength anteriorly and posteriorly and reported values 4 – 9 kPa below the norm. Tongue strength measured in the anterior position (56.50 ± 13.60 to 73.33 ± 12.03 kPa) was typically stronger than in the posterior position (52.00 ± 15.20 to 55.75 ± 13.58 kPa). In addition, findings from these three studies indicated that males (57.50 ± 15.10 to 73.33 ± 12.03 kPa) were stronger than females (56.50 ±13.60 to 61.27 ± 14.80 kPa) anteriorly but not posteriorly.

One study [62] reported values of tongue strength that were much lower than those reported by previous studies of healthy participants. Measures of tongue strength in this study were obtained while simultaneously recording from intramuscular electrodes inserted into the muscles of the tongue. No pre-electrode-insertion measures were obtained but one female participant was measured when only a few electrodes were inserted (value of 43 kPa) and again with all electrodes in place (29 kPa); a substantial decrease in tongue strength was observed with more electrodes, which

Table 3.3

Studies investigating tongue strength and endurance in healthy participants

					Tongue Strength (kPa) (mean ± SD)			Tongue Endurance (s) @ 50% Pmax (mean ± SD)		
Study name	**Year**	**Age range (y)**	***N***	**Sex**	**M & F across age groups**	**Males**	**Females**	**M & F across age groups**	**Males**	**Females**
ADULT STUDIES										
IOPI website										
young						**65.00**	**60.00**		**35.00**	**35.00**
old						**65.00**	**60.00**		**35.00**	**30.00**
Robin et al. * [63]	**1992**									
Trumpeters		18-48	12	8M, 4F	65.25 ± 11.74					
Control		18-49			65.98 ± 12.70					
Debaters		16-17	5	3M, 2F	77.63 ± 4.17					
Control		16-17			76.76 ± 6.00					
Robbins et al. * [49]	**1995**		24	24M						
young - blade		22-33				56.12				
- dorsum		67-83				48.02				
- tip						43.76				
old - blade						43.07				
- dorsum						39.32				
- tip						40.72				

Tongue Strength (kPa) (mean ± SD) **Tongue Endurance (s) @ 50% Pmax (mean ± SD)**

Study name	Year	Age range (y)	*N*	Sex	M & F across age groups	Males	Females	M & F across age groups	Males	Females
Crow & Ship * [38]	**1996**		99	52M, 47F		65.08 ± 18.90	56.29 ± 19.60			
		19-39			65.85 ± 17.30			43.90 ± 21.30		
		40-59			65.42 ± 23.60			41.90 ± 24.30		
		60-79			60.47 ± 17.30			48.00 ± 40.80		
		80-96			46.72 ± 13.30			45.20 ± 25.50		
Solomon et al. * [71]	**1996**	18-23	12	6M, 6F	60.47 ± 9.62	67.14 ± 9.13	53.80 ± 10.09			
Solomon et al. [66]	**2002**	19-26	10	5M, 5F	61.29 ± 8.80	65.82 ± 10.64	56.76 ± 6.45			
Solomon et al. [68]	**2004**	20-38	10	2M, 8F	61.60 ± 9.88		61.75 ± 9.53			
Youmans et al. [75]	**2006**		90	45M, 45F		64.00 ± 2.03	55.90 ± 1.86			
		20-39				72.00 ± 13.40	55.70 ± 12.50			
		40-59				63.90 ± 11.80	59.10 ± 14.00			
		60-96				56.10 ± 11.60	52.90 ± 10.70			
Palmer et al. [62]	**2008**		7	4M, 3F						
		24-37				49.25 ± 18.64				
		21-30					37.00 ± 11.36			

					Tongue Strength (kPa) (mean ± SD)			Tongue Endurance (s) @ 50% Pmax (mean ± SD)		
Study name	**Year**	**Age range (y)**	***N***	**Sex**	**M & F across age groups**	**Males**	**Females**	**M & F across age groups**	**Males**	**Females**

Study name	Year	Age range (y)	*N*	Sex	Tongue Strength (kPa) (mean ± SD): M & F across age groups	Tongue Strength (kPa) (mean ± SD): Males	Tongue Strength (kPa) (mean ± SD): Females	Tongue Endurance (s) @ 50% Pmax (mean ± SD): M & F across age groups	Tongue Endurance (s) @ 50% Pmax (mean ± SD): Males	Tongue Endurance (s) @ 50% Pmax (mean ± SD): Females
Vitorino et al. [73]	**2010**		75	35M, 40F	56.59 ± 2.73	56.81 ± 1.36	56.37 ± 4.07		15.72 ± 2.29	16.23 ± 2.11
young		20-40				58.18 ± 7.07	57.05 ± 8.48		15.12 ± 6.73	17.30 ± 10.03
middle		41-60				55.46 ± 7.69	60.06 ± 7.24		18.25 ± 7.32	17.60 ± 6.35
old		61-80				56.80 ± 6.87	52.00 ± 5.00		13.80 ± 2.05	13.80 ± 3.03
Kays et al. [8]	**2010**		22	10M, 12F						
young		20-35								
(anterior)						59.20 ± 5.20	67.80 ± 10.60		40.20 ± 14.00	37.50 ± 11.80
(posterior)						50.00 ± 7.90	62.50 ± 14.50		26.00 ± 19.50	29.60 ± 9.30
old		65-82								
(anterior)						62.60 ± 8.80	50.30 ± 11.10		29.60 ± 12.50	34.30 ± 19.30
(posterior)						61.40 ± 7.50	49.00 ± 12.60		24.20 ± 13.60	24.40 ± 14.40
Neel et al. [58]	**2011**	20-78	57	29M, 28F	65.28 ± 12.04	69.35 ± 10.85	61.00 ± 10.10		37.85 ± 9.69	25.45 ± 3.18
young males		20-40				74.10 ± 11.80			31.00 ± 17.40	
old males		22-40				64.80 ± 12.10			44.70 ± 28.40	
young females		42-78					64.60 ± 9.80			23.20 ± 9.20
old females		42-74					57.20 ± 7.60			27.70 ± 17.70
Gingrich et al. [54]	**2012**	18-34	30	15M, 15F						
(anterior)						73.33 ± 12.03	61.27 ± 14.80			
(posterior)						53.60 ± 14.33	50.07 ± 14.44			

Clark et al. [53]		**2012**	18-89	171	88M, 83F			
All males	(anterior)						57.50 ± 15.10	
	(posterior)						52.00 ± 15.20	
All females	(anterior)							56.50 ± 13.60
	(posterior)							53.60 ± 14.20
young	(anterior)		18-29			55.80 ± 13.50		
	(posterior)					52.30 ± 13.20		
middle	(anterior)		30-59			62.80 ± 13.00		
	(posterior)					57.90 ± 16.70		
old	(anterior)		60-89			51.00 ± 15.00		
	(posterior)					47.40 ± 16.70		
CHILD STUDIES								
Potter et al. (2009a) [81]			3-5	48	24M, 24F	28.50 ± 8.77		
Potter et al. (2009b) [82]			3-17	148	71M, 77F		48.08 ± 18.85	38.16 ± 8.14

* values in these studies multiplied by 0.87

explains the low values reported in this study. Males (range 34 to 72 kPa, mean 49 kPa) were again found to be stronger than females (range 32 to 50 kPa, mean 37 kPa).

Maximum tongue strength was observed to decrease with increasing age in nine studies involving healthy adults [49,75,73,37,58,8,38,83,54]. Results from these studies indicated that the oldest adults were, on average, 10 - 15 kPa lower than young adults. Two studies investigated tongue strength in healthy children. Potter et al. (2009a) studied children aged 3 - 5 years and found tongue strength increased with age ($p < 0.001$) [81]. Potter et al. (2009b) reported tongue strength in children and adolescents (3 - 17 years) and found significant differences in tongue strength with age up to10 years, after which no significant age-related differences were observed [82].

3.4.2 Tongue strength in populations with disorders

Seventeen studies (adults $n = 15$ and children $n = 2$) reported measures of tongue strength (in kPa) in populations with a disorder (Table 3.4). The main disorders were PD ($n = 5$), HNC ($n = 3$) and OPMD ($n = 2$). Mean values for PD ranged from 44.26 ± 3.22 kPa to 55.11 ± 13.82 kPa with higher tongue strength values in males than females. Three studies investigated HNC [55-57] with values ranging from 37.05 ± 14.42 kPa to 56.00 kPa. Lazarus et al. (2007) reported that mean maximum tongue strength was not significantly different to pre-treatment at one month after treatment but did increase significantly at 6- and 12- month post-treatment. Two studies investigated OPMD [61,59] with values much lower (19.50 ± 0.71 kPa to 26.90 ± 7.80 kPa) than healthy controls and those with other disorders such as PD.

Table 3.4

Studies investigating tongue strength and endurance in populations with a disorder

						Tongue strength (kPa) (mean ± SD)			Tongue endurance (s) (mean ± SD)		
Study name	**Year**	**Medical Condition**	**Age range (y)**	**N**	**Sex**	**M & F across age groups**	**Males**	**Females**	**M & F across age groups**	**Males**	**Females**
ADULT STUDIES											
Lazarus et al. * [55]	**2007**	HNC	29-78	46	35M, 11F						
(baseline)						47.00 ± 9.80					
(1 mth)						41.70 ± 8.22					
(3 mth)						51.00 ± 10.12					
(6 mth)						57.50 ± 10.12					
(12 mth)						54.70 ± 8.54					
Lazarus et al. [56]	**2000**	HNC	38-72	13	10M, 3F	37.05 ± 14.42			40.62 ± 24.67		
		Control	36-77	13	10M, 3F	60.15 ± 3.68			37.77 ± 3.18		
Lazarus et al. [57]	**2002**	HNC	72	1	1M		56.00			4.00	
		Control	72	1	1M		30.00			13.00	
Chang et al. [51]	**2008**	NPC	33-63	12	11M, 1F	56.67 ± 9.35			24.58 ± 10.72		
		Control	30-65	12	11M, 1F	64.50 ± 12.57			18.75 ± 6.22		
Neel et al. [59]	**2006**	OPMD	57-67	8	2M, 6F		19.50 ± 0.71	24.67 ± 9.09			
		Control	61,67	2	2F			50.50 ± 0.71			

						Tongue strength (kPa) (mean ± SD)			Tongue endurance (s) (mean ± SD)		
Study name	**Year**	**Medical Condition**	**Age range (y)**	**N**	**Sex**	**M & F across age groups**	**Males**	**Females**	**M & F across age groups**	**Males**	**Females**
Palmer et al. [61]	**2010**	OPMD	50-76	11	3M, 8F		26.90 ± 7.80	26.90 ± 7.80			
		Control	52-76	9	4M, 5F		57.40 ± 10.40	57.40 ± 10.40			
Solomon et al. [69]	**1994**	PD	43-71	3	1M, 2F		53.00	49.50		6.00	50.00
		Control	43-64	3	1M, 2F		70.00	51.50		25.00	37.50
Solomon et al. [67]	**1995**	PD	46-73	19	10M, 9F		52.98 ± 19.93	50.07 ± 16.79		23.23 ± 11.14	34.32 ± 47.69
		Control	49-74	19	10M, 9F		63.25 ± 10.66	56.94 ± 9.68		23.14 ± 11.58	28.90 ± 11.44
Solomon et al. [70]	**2000**	PD	56-81	16	12M, 4F		48.25 ± 10.04	47.75 ± 10.21		21.10 ± 9.52	22.20 ± 20.81
		Control	55-93	16	12M, 4F		53.75 ± 6.18	60.75 ± 14.95		38.46 ± 32.05	32.05 ± 9.84
Solomon [64]	**2006**	PD	40-75	12	9M, 3F		55.11 ± 13.82	49.00 ± 20.42			
		Control	48-74	15	8M, 7F		63.75 ± 13.96	57.00 ± 7.59			
Robin et al. [84]	**1991**	TBI	26	1	1F			38.28			25.00
		Control	20-49	26	5M, 21F	61.77					36.31 ± 10.13
Yeates et al. * [74]	**2008**	TBI, HNC,	50-72	3	3M						
(anterior)		CVA					45.25 ± 19.37				
(posterior)							42.24 ± 21.95				
Clark et al. * [6]	**2003**	Various	19-95	63	28M, 35F	32.75 ± 18.44					

Tongue strength (kPa) (mean ± SD)	Tongue endurance (s) (mean ± SD)

Study name	Year	Medical Condition	Age range (y)	N	Sex	M & F across age groups	Males	Females	M & F across age groups	Males	Females
Solomon et al. * [65]	**2008**	Various	18-78	44	40M, 4F		43.18 ± 20.00	48.25 ± 13.82		38.28 ± 24.57	32.00 ± 21.83
Stierwalt & Youmans [72]	**2007**	Various	26-91	50	16M, 26F		42.89 ± 15.60	31.03 ± 15.85		49.85 ± 52.27	37.77 ± 37.30
		Control	26-90	50	16M, 26F		63.24 ± 13.86	57.15 ± 13.50		42.77 ± 16.14	37.15 ± 30.55
CHILD STUDIES											
Robin et al. [84]	**1991**	DAS, DVD	8-10	5	4M, 1F	40.02			9.10 ± 4.84		
		Control	6-12	6	4M, 2F	56.55			24.03 ± 4.13		
Stierwalt et al. [85]	**1996**	TBI	6-17	23	14M, 9F		56.24 ± 18.67	36.05 ± 13.58		14.50 ± 14.47	8.78 ± 10.54
		Control	6-17	23	14M, 9F		64.44 ± 11.82	47.56 ± 9.73		38.14 ± 17.10	24.00 ± 19.91

* No control group used

HNC=Head or neck cancer; NPC=Nasopharyngeal cancer; OPMD=Oculopharyngeal muscular dystrophy; PD=Parkinson's Disease; TBI=Traumatic brain injury; DAS=Developmental apraxia of speech; DVD=Developmental verbal dyspraxia

3.4.3 Tongue endurance in healthy populations

Tongue endurance (reported in seconds) was measured isometrically at 50% of maximal tongue strength (P_{max}) in the anterior position (unless otherwise stated) and reported in four studies (Table 3.3) in healthy people. Effects of age on tongue endurance in males and females in four age groups (young, middle-aged, older, and elderly) was examined [38]. Regardless of age or sex, overall mean tongue endurance was 44.80 ± 28.00 s, and no significant differences in tongue endurance with age were observed (p = 0.67). Mean tongue endurance values ranged from 15.72 ± 5.86 to 37.85 ± 23.55 s for males and 16.23 ± 7.07 to 36.35 ± 11.74 s for females, with no significant age effects in either males (p = 0.61) or females (p = 0.33). A comparison of tongue endurance in two age groups (20 - 35y and 65 - 82y) and in two positions on the tongue (anterior and posterior) was conducted [8]. Significant differences in tongue endurance were observed in the anterior compared to the posterior position (p = 0.0005) but no significant age or sex differences were reported. Neel et al. (2011) examined tongue endurance in males and females in two age groups (20 - 40y and 42 - 78y). Males had higher values than females (p < 0.03) and there was a trend for older adults to have higher values than younger adults (p < 0.10). The mean values for each subgroup were older males (44.70 ± 28.40 s), younger males (31.00 ± 17.40 s), older females (27.70 ± 17.70 s) and younger females (23.20 ± 9.20 s). Vitorino et al. (2010) examined three age groups (20 - 40y; 41 - 60y; and 61 - 80y), and their tongue endurance measures were lower than those in other studies, however no significant differences were reported across age (p > 0.05) or sex (p > 0.05). Robin et al. (1992) investigated tongue endurance in individuals with high skills levels with their tongues (trumpet players and debaters). Although values were not provided (other than in a figure), they reported that both debaters and trumpet players had substantially higher values than healthy controls.

3.4.4 Tongue endurance in populations with disorders

Ten studies (adults *n* = 9 and children *n* = 1) measured tongue endurance (in seconds) isometrically at 50% of maximum tongue strength in populations with disorders (Table 3.4). Five disorders accounted for most of those measured: PD, HNC; OPMD; NPC; TBI. Three studies measured endurance with values ranging from 6.00 to 23.23 ± 11.14 s compared to a control group (23.14 ± 11.58 to 38.46 ± 32.05 s). Females in PD studies (22.20 ± 20.81 s) were better able to hold 50% maximum tongue strength than males (21.10 ± 9.52 s). Stierwalt and Youmans (2007) examined various medical conditions including 29 participants following CVA with males reporting longer endurance times (49.85 ± 52.27 s) than females (37.77 ± 37.30 s) [86]. No endurance data was available for individuals following CVA. One study [84] investigated children (DAS, DVD) with males (14.50 ± 14.47 s) having better endurance than the females (8.78 ± 10.54 s). Males in the control group (38.14 ± 17.10 s) also had longer endurance times than female controls (24.00 ± 19.91 s). This study also reported that children with DVD and/or DAS (9.10 ± 4.84 s) were not able to hold an endurance level similar to the control group (24.03 ± 4.13 s) [84].

Comparisons with healthy control groups indicate that populations with disorders have significantly lower tongue endurance, with the magnitude of the decrease dependent on the specific medical condition; this is demonstrated in a study examining OPMD in older adults by Palmer et al. (2010). Compared to a control group, the OPMD group showed a decrease in tongue endurance however it was not significant [61].

3.4.5 Hand strength in healthy populations

Only three studies (adults *n* = 2 and children *n* = 1) reported hand strength (kPa) in healthy individuals (Table 3.5). Such a small number of studies provides little basis for the establishment of normative hand strength values in healthy adults. Crow and Ship (1996) investigated the effects of age and sex in healthy adults with males (155.10 ± 44.60 kPa) stronger ($p<0.001$) than females (123.60 ± 27.20 kPa). Younger adults had the highest values (165.00 ± 43.80 kPa), followed by middle-aged (157.70 ± 34.10 kPa), older (129.00 ± 35.30 kPa), and elderly (110.00 ± 33.20 kPa) groups. Mean hand strength across broader age groups was also reported (140.43 ± 36.60 kPa) with a significant difference in strength ($p < 0.01$) between individuals aged greater than 59 years and younger age groups. Robin et al. (1992) reported hand strength values for trumpet players (157.34 ± 25.74 kPa) and a control group (171.58 ± 23.32 kPa) with significance observed ($p < 0.0001$). A debaters group (171.35 ± 13.20 kPa) showed values that were also significant ($p < 0.0002$) when compared to a control group (181.13 ± 23.32 kPa). Potter et al. (2009a) reported mean hand strength of 48.41 ± 8.18 kPa in 48 children aged 3 to 5 years [81].

Table 3.5

Studies investigating hand strength and endurance in a healthy population

Study name	Year	Age range (y)	*N*	Sex	Hand Strength (kPa) (mean ± SD) M & F across age groups	Hand Strength Males	Hand Strength Females	Hand Endurance (s) (mean ± SD) M & F across age groups	Hand Endurance Males	Hand Endurance Females
ADULT STUDIES										
IOPI website						**150.00**	**140.00**		**40.00 - 60.00**	**40.00 - 60.00**
Robin et al. [63]	**1992**									
Trumpeters		18-48	12	8M, 4F	157.35 ± 25.74					
Control		18-49			171.58 ± 23.32					
Debaters		16-17	5	3M, 2F	171.35 ± 13.20					
Control		16-17			181.13 ± 23.32					
Crow et al. [38]	**1996**	19-96	99	52M, 47F		155.10 ± 44.60	123.60 ± 27.20		74.20 ± 38.30	90.30 ± 49.80
		19-39			165.00 ± 43.80			72.30 ± 44.30		
		40-59			157.70 ± 34.10			88.50 ± 39.60		
		60-79			139.00 ± 35.30			84.20 ± 46.60		
		80-96			110.00 ± 33.20			72.60 ± 50.50		
CHILD STUDY										
Potter et al. [81]	**2009a**	3-5	48	24M, 24F	48.41 ± 8.18					

3.4.6 Hand strength in populations with disorders

Five studies (adults) reported measures of hand strength (in kPa) in populations with medical conditions (Table 3.6), primarily PD. Two studies [67,70] examined hand strength in older adults with PD. Solomon et al. (1995) reported that male values (131.20 ± 29.84 kPa) were stronger than females (94.83 ± 35.36 kPa) but not as strong as the age and sex matched control groups (males 150.08 ± 34.13 and females 120.64 ± 25.16). Solomon et al. (2000) also reported values for males and females with PD (140.33 ± 23.46 kPa and 98.25 ±14.31 kPa respectively), however these were not significantly different ($p = 0.362$) to male and female control group participants (136.58 ± 23.75 kPa and 101.75 ± 24.88 kPa respectively).

3.4.7 Hand endurance in healthy populations

Two studies (adults $n = 1$ and children $n = 1$) measured hand endurance in seconds at 50% of maximum hand strength. One study [38] measured hand endurance in healthy adults (Table 3.5). Mean hand endurance regardless of age was 79.40 ± 45.25 s, and there were no significant differences in hand endurance with age whether analysed with all participants ($p = 0.41$), or for males ($p = 0.38$) or females ($p = 0.56$). Mean values reported for different age groups were middle-aged adults (88.50 ± 39.60 s), adults (84.20 ± 46.60 s), elderly adults (72.60 ± 50.50 s) and younger adults (72.30 ± 44.30 s). There was a trend ($p = 0.08$) for females to sustain hand endurance longer (90.30 ± 49.80 s) than males (74.20 ± 38.30 s). Robin et al. (1991) examined hand endurance in 26 healthy adults and six healthy children. Children sustained hand endurance for an average of 24.03 ± 4.13 s while adults averaged 36.31 ± 10.13 s ($p < 0.05$).

Table 3.6

Studies investigating hand strength and endurance in populations with a disorder

Study name	Medical Condition	Age range (y)	*N*	Sex	Hand strength (kPa) (Mean ± SD)			Hand Endurance (s) (Mean ± SD)	
					M & F across ages	Males	Females	Males	Females
ADULT STUDIES									
Robin et al. (1991) [84]	TBI	26	1	1F			132.00		56.00
	Control	20-49	26	5M, 21F	110.00				56.49 ± 13.70
Solomon et al. (1994) [69]	PD	43-71	3	1M, 2F		273.00	131.75	33.00	67.50
	Control					156.00	147.50	24.00	45.00
Solomon et al. (1995) [67]	PD	46-72	19	10M, 9F		131.20 ± 29.84	94.83 ± 35.36	44.81 ± 45.95	46.50 ± 18.48
	Control					150.08 ± 34.13	120.64 ± 25.16	41.67 ± 21.98	48.72 ± 20.24
Solomon et al. (2000) [70]	PD	56-81	16	12M, 4F		140.33 ± 23.46	98.25 ± 14.31	53.18 ± 20.79	63.40 ± 39.48
	Control					136.58 ± 23.75	101.75 ± 24.88	57.38 ± 16.19	60.63 ± 50.63
O'Day et al. (2005) [60]	PD	52 – 79	10	10M					
	day 1					105.90 ± 32.93			
	day 2					106.10 ± 28.93			
	day 3					110.50 ± 38.55			
	day 4					109.20 ± 31.62			
	day 5					111.70 ± 38.67			
	Control								
	day 1					133.20 ± 25.62			
	day 2					139.30 ± 25.27			
	day 3					136.90 ± 24.03			
	day 4					134.20 ± 23.71			
	day 5					137.50 ± 18.09			

Study name	Medical Condition	Age range (y)	*N*	Sex	Hand strength (kPa) (Mean ± SD)			Hand Endurance (s) (Mean ± SD)	
					M & F across ages	Males	Females	Males	Females
CHILD STUDY									
Robin et al. (1991) [84]	DAS, DVD	8-10	5	4M, 1F					11.57 ± 6.96
	Control	6-12	6	4M, 2F					48.00 ± 10.14

HNC=Head or neck cancer; NPC=Nasopharyngeal cancer; OPMD=Oculopharyngeal muscular dystrophy; PD=Parkinson's Disease; TBI=Traumatic brain injury;

DAS=Developmental apraxia of speech; DVD=Developmental verbal dyspraxia

3.4.8 Hand endurance in populations with disorders

Five studies (adults n = 4 and children n = 1) reported measures of hand endurance in populations with disorders (Table 3.6). Three studies examined PD [67,69,70]. Solomon et al (1994) reported three case studies (one male and two females) and found reduced or abnormal findings for hand endurance. Solomon et al. (1995) reported values for males (44.81 ± 45.95 s) and females (46.50 ± 18.48 s) with a statistically significant difference between PD and control groups ($p = 0.025$). Solomon et al. (2000) reported values for males (53.18 ± 20.79 s) and females (63.40 ± 39.48 s) with no significant difference between the disordered and control groups ($p = 0.805$). Stierwalt et al. (1996) measured hand endurance in 23 children with TBI compared to a control group and found a significant difference between groups ($p = 0.0001$) [85]. One study [84] reported a value of 11.57 ± 6.96 s for children aged 8 to 10 years with DAS, which was significantly different ($p < 0.05$) to the healthy control group (48.00 ± 10.14 s). This study also reported values for one female with TBI (56.00s) and found a comparable result to a control group (56.49 ± 13.70s) (no p value reported).

3.5 Results of meta-analyses

Meta-analyses were conducted for tongue strength and endurance for age and sex. Funnel plot comparison for meta-analyses 2, 3 and 4 were not generated as less than 10 studies were included. Meta-analysis was deemed inappropriate for younger participants (< 60 years) vs. older participants (60+ years) for males and females as results from fewer than three studies were compatible for analysis.

3.5.1 Tongue strength : Meta-analysis 1

In total, males (n=425) and females (n=391) (total 816) from 17 studies with ages ranging from 19 to 96 years were included. The studies were statistically heterogeneous ($Tau^2 = 20.05$; $\chi^2 = 112.78$, $df = 16$, $P < 0.00001$, $I^2 = 86\%$), so the random effects model was used.

Meta-analysis (Figure 3.2) revealed statistically significant greater tongue strength in males compared to females (WMD 5.21kPa [2.26, 8.17; 95% CI], Z = 3.46, p = 0.0005). As this meta-analysis used a random-effect estimate funnel plot comparison for tongue strength to assess publication bias was not generated even though greater than 10 studies were included. Random-effects estimates give greater relative weight to smaller studies and may lead to wider CIs [79].

Study or Subgroup	Males Mean	Males SD	Males Total	Females Mean	Females SD	Females Total	Weight	Mean Difference IV, Random, 95% CI
Youmans [2]	0	2.03	45	55.9	1.86	0		Not estimable
Solomon [13]	61	15.56	2	61.75	9.53	8	1.1%	-0.75 [-23.30, 21.80]
Palmer [14]	49.25	18.64	4	37	11.36	3	1.2%	12.25 [-10.09, 34.59]
Solomon [19]	53.75	6.18	12	60.75	14.95	4	2.4%	-7.00 [-22.06, 8.06]
Solomon [28]	43.18	20	40	48.25	13.82	4	2.4%	-5.07 [-19.96, 9.82]
Solomon [24]	63.75	13.96	8	57	7.59	7	3.9%	6.75 [-4.44, 17.94]
Solomon [12]	65.82	10.67	5	56.76	6.45	5	4.1%	9.06 [-1.87, 19.99]
Solomon [11]	67.14	9.13	6	53.8	10.09	6	4.1%	13.34 [2.45, 24.23]
Gingrich [7]	73.33	12.03	15	61.27	14.8	15	4.9%	12.06 [2.41, 21.71]
Solomon [17]	63.25	10.66	10	56.94	9.68	9	5.3%	6.31 [-2.84, 15.46]
Kays [15]	60.9	7.046	10	59.05	13.8058	12	5.5%	1.85 [-7.10, 10.80]
Stierwalt [26]	63.24	13.86	16	57.15	13.5	26	5.8%	6.09 [-2.46, 14.64]
Crow [10]	65.08	18.9	52	56.29	19.6	47	6.8%	8.79 [1.19, 16.39]
Neel [6]	69.931	12.6291	29	60.5172	9.2806	29	9.3%	9.41 [3.71, 15.12]
Clark [43]	57.5	15.1	88	56.5	13.6	83	11.8%	1.00 [-3.30, 5.30]
Youmans [3]	69.99	7.49	48	66.96	8.29	48	14.1%	3.03 [-0.13, 6.19]
Vitorino [4]	56.81	1.36	35	56.37	4.07	40	17.3%	0.44 [-0.90, 1.78]
Total (95% CI)			**425**			**346**	**100.0%**	**4.50 [2.00, 6.99]**

Heterogeneity: Tau² = 8.91; Chi² = 30.78, df = 15 (P = 0.009); I² = 51%
Test for overall effect: Z = 3.53 (P = 0.0004)

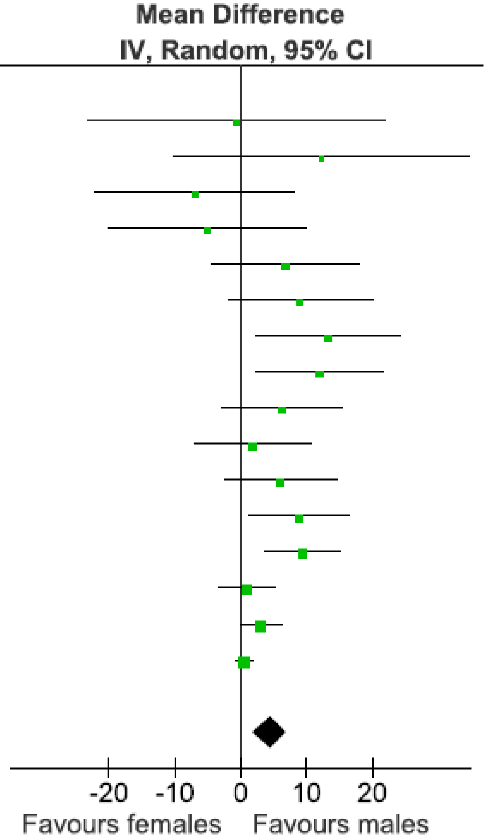

Figure 3.2. Forest plot of comparison: Tongue Strength by Age and Sex, Males vs. Females.

Note. References listed above in square brackets [ ] relate to the publication provided in the systematic review published in 2013.

3.5.2 Tongue strength : Meta-analysis 2

Two age groups were considered: (< 60 years = younger and 60+ years = older). Data from adults less than 60 years (*n* = 484) were compared to adults 60+ years (*n* = 275) (total 759) from eight studies. The studies were not statistically heterogeneous (χ^2 = 3.54, *df* = 7, *p* = 0.83, I^2 = 0%), so the fixed effects model was used. Meta-analysis (Figure 3.3) revealed statistically significant greater tongue strength in adults less than 60 years compared to adults 60+ years (WMD 8.30 kPa [6.37, 10.23], Z = 8.43 (P < 0.00001).

3.5.3 Tongue strength : Meta-analysis 3

Two age groups were considered (< 60 years = younger and 60+ years = older) for males. In total, younger males (*n* = 93) vs. older males (*n* = 63) (total 156) from five studies were included. Studies were not statistically heterogeneous (χ^2 = 7.83, *df* = 4 (P = 0.10); I^2 = 49%), so the fixed effects model was used. Meta-analysis (Figure 3.4) revealed that younger males had significantly stronger tongue strength than older males (WMD 8.00 kPa [4.92, 11.08; 95% CI], Z = 5.09 (P < 0.00001).

3.5.4 Tongue strength : Meta-analysis 4

Two age groups were considered (< 60 years = younger and 60+ years = older) for females. In total, younger females (*n* = 80) vs. older females (*n* = 53) (total 133) from four studies were included. Studies were not statistically heterogeneous (χ^2 = 5.40, *df* = 3 (P = 0.14); I^2 = 44%), so the fixed effects model was used. Meta-analysis (Figure 3.5) revealed that younger females had significantly stronger tongue strength than older females (WMD 9.43 kPa [5.57, 13.28; 95% CI], Z = 4.79 (P < 0.00001).

	Adults < 60y			Adults 60+ y				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI
Clark [43]	59.0812	13.6732	128	51	15	43	14.5%	8.08 [3.01, 13.15]
Crow [10]	65.58	21.2537	43	57.278	17.3625	56	6.1%	8.30 [0.49, 16.11]
Kays [15]	63.8909	9.3363	11	55.8909	11.569	11	4.8%	8.00 [-0.79, 16.79]
Neel [6]	70.0286	11.8069	28	60.6069	10.4156	29	11.1%	9.42 [3.63, 15.21]
Robbins [9]	49.3	7.21	10	41.04	2.17	14	17.5%	8.26 [3.65, 12.87]
Stierwalt [26]	61.5057	13.3688	140	55.01	14.32	60	20.6%	6.50 [2.25, 10.74]
Youmans [2]	62.675	13.5599	60	56.1	11.6	30	12.8%	6.57 [1.19, 11.96]
Youmans [3]	72.6475	12.7437	64	60.125	12.9415	32	12.5%	12.52 [7.06, 17.99]
Total (95% CI)			**484**			**275**	**100.0%**	**8.30 [6.37, 10.23]**

Heterogeneity: Chi² = 3.54, df = 7 (P = 0.83); I² = 0%
Test for overall effect: Z = 8.43 (P < 0.00001)

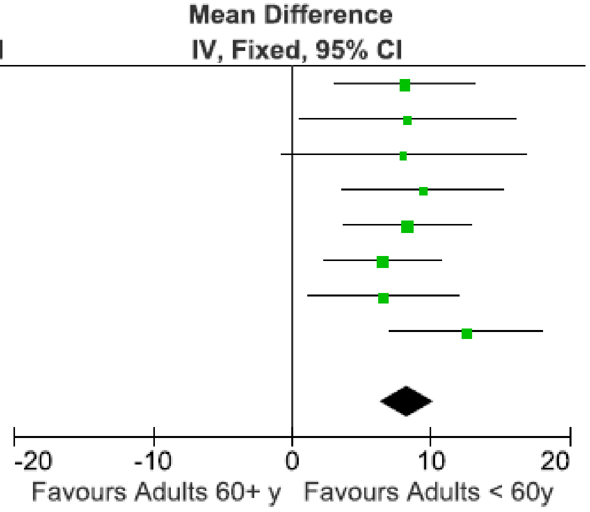

Figure 3.3. Forest plot of comparison: Tongue Strength by Age and Sex, Adults < 60y vs. Adults 60+ y.

Note. References listed above in square brackets [ ] relate to the publication provided in the systematic review published in 2013.

	Younger males < 60y			Older males 60+y				Mean Difference
Study or Subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 95% CI**
Kays [15]	59.2	5.2	5	62.6	8.8	5	11.8%	-3.40 [-12.36, 5.56]
Neel [6]	74.1	11.8	16	64.8	12.1	13	12.4%	9.30 [0.54, 18.06]
Robbins [9]	49.3	7.21	10	41.04	2.17	14	44.6%	8.26 [3.65, 12.87]
Youmans [2]	67.95	13.0718	30	56.1	11.6	15	16.8%	11.85 [4.34, 19.36]
Youmans [3]	73.64	14.444	32	62.69	13.04	16	14.4%	10.95 [2.83, 19.07]
Total (95% CI)			**93**			**63**	**100.0%**	**8.00 [4.92, 11.08]**

Heterogeneity: Chi² = 7.83, df = 4 (P = 0.10); I² = 49%
Test for overall effect: Z = 5.09 (P < 0.00001)

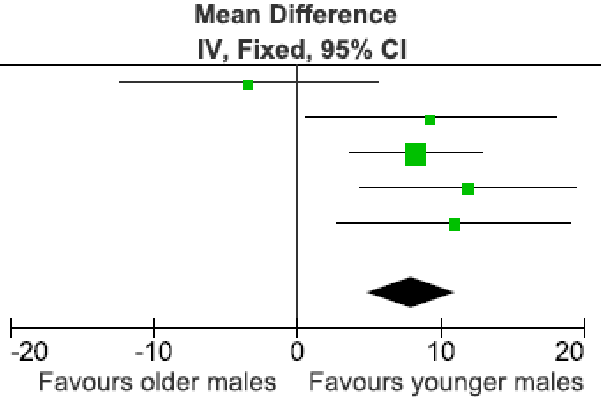

Figure 3.4. Forest plot of comparison: Tongue Strength by Age and Sex, Younger males vs. Older males.

Note. References listed above in square brackets [ ] relate to the publication provided in the systematic review published in 2013.

	Younger females < 60y			Older females 60+y				Mean Difference
Study or Subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 95% CI**
Kays [15]	67.8	10.6	6	50.3	11.1	6	9.9%	17.50 [5.22, 29.78]
Neel [6]	64.6	9.8	12	57.2	7.6	16	33.3%	7.40 [0.72, 14.08]
Youmans [2]	57.4	13.1545	30	52.9	10.7	15	28.9%	4.50 [-2.67, 11.67]
Youmans [3]	71.655	10.926	32	57.56	12.73	16	27.9%	14.09 [6.80, 21.39]
Total (95% CI)			**80**			**53**	**100.0%**	**9.43 [5.57, 13.28]**

Heterogeneity: Chi² = 5.40, df = 3 (P = 0.14); I² = 44%
Test for overall effect: Z = 4.79 (P < 0.00001)

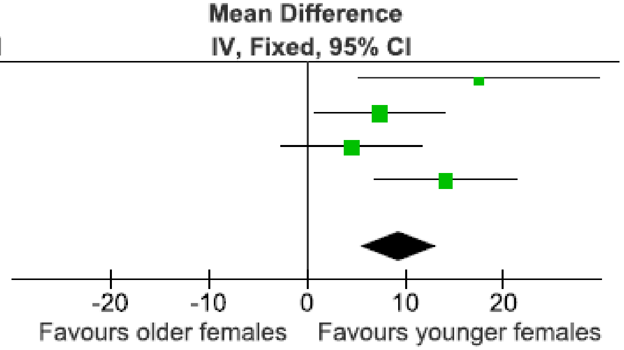

Figure 3.5. Forest plot of comparison: Tongue Strength by Age and Sex, Younger females vs. Older Females.

Note. References listed above in square brackets [ ] relate to the publication provided in the systematic review published in 2013.

3.5.5 Tongue endurance : Meta-analysis 1

One meta-analysis was conducted with 112 males and 119 females (total 231) from six studies included. The evaluation studies were statistically heterogeneous ($\chi^2 = 7.37$, $df = 5$, $p = 0.19$, $I^2 = 32\%$), so the fixed effects model was used. Meta-analysis (Figure 3.6) revealed no statistically significant difference in tongue endurance between adult males and females across all ages (WMD -0.40 seconds [-1.39, 0.58; 95% CI], Z = 0.80, $p = 0.42$).

3.6 Intervention studies

Five studies investigated the effects of interventions on the strength and endurance of the tongue (Table 3.7). Two RCTs [35,87] used the IOPI as both an intervention and evaluation tool, and evaluated the effects of tongue-strengthening exercises on tongue strength and endurance in healthy adults. The third RCT [88] randomised participants to five tongue training groups (strength, endurance, power, speed, and no training) and used the IOPI for the measurement of tongue strength, endurance and power, but not speed, pre- and post-training. Participants in the two prospective cohorts studies [17,18] used the IOPI to measure tongue strength and endurance following an eight-week tongue-strengthening exercise program in older-adult healthy and stroke populations.

Studies varied in the following areas: age groups (18 - 67y, 19 - 57y; 20 - 29y, 51 - 90y, 70 - 89y), medical condition (healthy, stroke); sex imbalance (more females than males); study duration (4, 8 or 9 weeks); participant group size (10, 31, 31 and 39); frequency of measurements (time series, fortnightly, or monthly); exercise program (10 repetitions 3 times/day on 3 non-consecutive days; 10 repetitions 3 times/day for 7 days/week; 10 repetitions 5 times/day for 5 days/week; or 3 sessions per week on 3 non-consecutive days for 4 weeks). Outcome measures (tongue strength and

	Males			Females				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Fixed, 95% CI
Kays [15]	34.9	13.7028	10	35.9	15.3427	12	0.7%	-1.00 [-13.14, 11.14]
Neel [6]	37.1414	23.5784	29	25.7714	14.6177	28	0.9%	11.37 [1.22, 21.52]
Solomon [17]	23.14	11.58	10	28.9	11.44	9	0.9%	-5.76 [-16.12, 4.60]
Solomon [19]	38.46	32.05	12	32.05	9.84	4	0.2%	6.41 [-14.13, 26.95]
Stierwalt [26]	42.77	16.14	16	37.15	30.55	26	0.5%	5.62 [-8.54, 19.78]
Vitorino [4]	15.72	2.29	35	16.23	2.11	40	96.8%	-0.51 [-1.51, 0.49]
Total (95% CI)			**112**			**119**	**100.0%**	**-0.40 [-1.39, 0.58]**

Heterogeneity: $Chi^2 = 7.37$, $df = 5$ ($P = 0.19$); $I^2 = 32\%$

Test for overall effect: $Z = 0.80$ ($P = 0.42$)

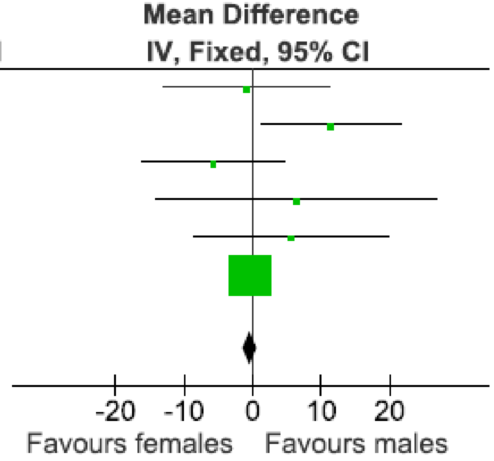

Figure 3.6. Forest plot of comparison: Tongue Endurance by Age and Sex, Males vs. Females.

Note. References listed above in square brackets [ ] relate to the publication provided in the systematic review published in 2013.

endurance; only tongue strength; or tongue strength and endurance within specific training groups); tongue bulb position (anterior only, or anterior and posterior); and training specificity (directional exercise – elevation, protrusion, lateralisation, or none) were reported. The RCT by Clark (2012) differed from the other four intervention studies in that it reported Cohen's *d* values as well as *p* values.

3.6.1 Tongue strength

Four studies examined tongue strength pre- and post-tongue-strengthening exercise programs (Table 3.7). Lazarus et al. (2003) investigated the effects of IOPI or tongue depressor exercise training in young adults (20 - 29y); the responses of the two exercise intervention groups did not differ and when combined showed significant improvements from baseline (64.40 ± 8.71 kPa) to four weeks (73.10 ± 7.33 kPa) compared to a no-exercise control group ($p = 0.04$). Robbins et al. (2005) examined the effects of six weeks IOPI exercise training in older adults (70 - 89y). Significant increases in tongue strength were observed from baseline to four weeks ($p = 0.002$) and baseline to six weeks ($p = 0.001$), with the following values (in kPa) reported: baseline (41.00; range 36 - 46); two weeks (44.00; range 39 - 49); four weeks (47.00; range 43 - 51); and six weeks (49.00; range 45 - 53). Clark et al. (2009) examined the effects of nine weeks of training using three different directional exercise conditions (elevation, protrusion and lateralization) on tongue strength measured with the IOPI in healthy adults (18 - 67y). Training effects were reported at three and nine weeks. Significant increases in strength were observed with a 6% change in elevation strength ($p < 0.001$) compared to 26.6% for lateralisation ($p < 0.001$) and 13.4% for protrusion ($p <0.001$). Clark (2012) examined the specificity of exercise training effects using the IOPI in healthy adults (19 - 57y). Large ($d = 1.06$) improvements in strength were observed for the strength-training group only.

3.6.2 Tongue endurance

Two intervention studies investigated the effects of exercise training on tongue endurance (Table 3.7). Lazarus et al. (2003) examined the effects of IOPI or tongue depressor exercise training on tongue endurance in young adults (20 - 29y), and showed a trend to improve from baseline (25.00 ± 14.21 s) to four weeks (34.40 ± 31.62 s) ($p = 0.10$). Dosage included 10 repetitions completed five times per day for five days per week for four weeks with each repetition held for two seconds and performed in four directions (i.e. left, right, on protrusion, and on elevation.

Clark (2012) assessed tongue endurance using the IOPI to determine the effects of four different types of exercise training, including elevation exercises, which required the tongue to be pressed against the hard palate just behind the alveolar ridge with maximum effort. Dosage included 30 repetitions for 7 days per week in sets of 10 repetitions for three sets of elevation, protrusion and lateralisation. Clark (2012) found that endurance training had a large effect ($d = 1.29$) on isotonic tongue endurance (repetitions) but no effect on isometric endurance.

Table 3.7

Studies investigating the use of the IOPI in intervention studies

							Tongue strength (kPa) (mean ± SD)		Tongue endurance (s) (mean ± SD)	
Study name	**Study Design**	**Groups**	**Health Status**	**Age groups (y)**	***n***	**Sex**	**Baseline**	**Post-exercise**	**Baseline**	**Post-exercise**
ADULT STUDIES										
Robbins et al. (2005) [17]	Prospective cohort	IOPI	Healthy	70-89	10	4M, 6F	41.00	49.00		
Clark et al. (2009) [35]	RCT		Healthy	18-67	39	17M, 22F				
		TD Elevation					59.63 ± 14.12			
		TD Protrusion						66.65 ± 14.50		
		TD Lateral						66.46 ± 14.13		
								66.45 ± 14.91		
Lazarus et al. (2003) [87]	RCT		Healthy	20-29	31	12M, 23F				
		TD					64.80 ± 9.48	74.00 ± 7.59	29.70 ± 16.44	43.70 ± 43.96
		IOPI					63.90 ± 6.96	72.10 ± 6.64	20.80 ± 10.75	26.00 ± 9.49
		IOPI & TD					64.40 ± 8.71	73.10 ± 7.33	25.00 ± 14.21	34.40 ± 31.62
		Control					69.80 ± 17.71	71.20 ± 17.08	17.90 ± 8.22	18.40 ± 8.54
Clark (2012) [88]	RCT	IOPI	Healthy	19-57	25	3M, 22F				
		Strength trg					65.80 ± 14.97	82.60 ± 13.39	45.20 ± 10.28	45.40 ± 10.16
		Endurance trg					65.60 ± 15.19	73.00 ± 18.40	81.20 ± 32.41	77.20 ± 12.09
		Power trg					60.20 ± 17.98	66.60 ± 17.05	71.60 ± 30.22	71.40 ± 12.58
		Speed trg					72.80 ± 14.72	80.40 ± 20.11	62.80 ± 9.07	64.40 ± 12.70
		Control					66.80 ± 13.18	73.60 ± 10.06	62.40 ± 5.18	59.80 ± 14.48
Robbins et al. (2007) [18]	Prospective cohort	IOPI	Stroke	51-90	10	5M, 5F				
		Anterior					35.6	51.7		
		Posterior					30.2	54.6		

TD = Tongue depressor: RCT = randomised control trial

3.6.3 Risk of bias in intervention studies

The results of the 10-item risk of bias analysis for the five intervention studies are presented in Table 3.8. Inter-rater reliability for the risk of bias items between two reviewers (VA and RC) indicated a high level of agreement across all items (percentage agreement 100%, Cohen's $\kappa = 1$). Each study received a point for each indicator that met the quality criteria. For the three randomised controlled trials, all eight-quality indicators were relevant, leading to a maximum quality score of 8. For the other study designs, where an intention-to-treat analysis was not applicable, the highest quality score was seven. Randomisation was described adequately and performed in two studies [35,88] and a control group (randomised participants) was included in two studies [88,35,87]. Assessor blinding was carried out in only one study [35]; baseline characteristics were reported and at least one primary outcome measure was valid and reliable in all five studies [88,35,87,17,18]; *p* values were reported in five studies [35,87,17,18]. Effect sizes and/or precision estimates (e.g., 95% CIs) were reported in two studies [18]; magnitude of effect size was determined in two studies [88] using Cohen's benchmarks for small, medium, and large effects as 0.2, 0.5, and 0.8, respectively [89]. Summary results for individual study groups were presented in all studies cited. One study [35] indicated a low risk of bias with six of the eight quality markers. Four studies [87,88,18] and Robbins et al. (2005) had a higher risk of bias for four and three quality markers respectively.

Table 3.8

Risk of bias assessment of intervention studies

Studies	Did the study include a true control group (randomised participants not a comparison group)?	Were the assessors blinded to treatment allocation at baseline and post-test?	Was the randomisation procedure adequately described and carried out?	Were the subjects at baseline adequately described?	Was at least one primary outcome measure valid and reliable?	Did the study report or calculate a *p* value?	Did the study report effect size or confidence intervals?	Did the study report a power calculation and was the study adequately powered to detect intervention effects?
Clark et al. (2009) [35]	1	1	1	1	1	1	0	0
Clark (2012) [88]	1	0	1	1	1	1	1	1
Lazarus et al. (2003) [87]	1	0	0	1	1	1	0	0
Robbins et al. (2007) [18]	0	0	0	1	1	1	1	0
Robbins et al. (2005) [17]	0	0	0	1	1	1	0	0

1 = yes; 0 = no

Score of 0 - 5 = high risk of bias; score of 6 - 8 = low risk of bias

3.7 Discussion

This review systematically examined the state and quality of the evidence for the use of the IOPI to measure strength and endurance of the tongue and hand in healthy populations and those with a range of medical conditions. A systematic search of the scientific literature published since 1991 yielded 38 studies that addressed this purpose. The IOPI was used mostly for tongue strength (38 studies) and endurance (15 studies) measurement; relatively few studies measured hand strength (9 studies) or endurance (6 studies). Most of the studies used the IOPI as an evaluation tool, although four studies also used it as an intervention tool. Half the studies were conducted in healthy people, mostly in adults. Most of the other participants had disorders associated with dysphagia, such as PD or HNC. In healthy populations, both age and sex influence the tongue strength values obtained, but there is no sex difference in tongue endurance values.

3.8 Consolidation of results

3.8.1 Tongue strength

The IOPI has been most widely used to measure tongue strength, which was the rationale for its original development [84]. Tongue strength can be measured in different tongue positions, and anterior measurements produce higher values than posterior measurements. Measures of tongue strength taken in the anterior position showed that males typically generate higher values than females, but this difference appears to be absent or substantially reduced when posterior measurements are used [54,8,53]. Issues about where the bulb is in the mouth on recording tongue strength are important to note because of the possibility of slippage in the anterior and posterior positions. The average discrepancy between male and female values of tongue strength in healthy populations was 5.2 kPa, as suggested previously [7].

Age also influenced the values obtained, with strength increasing with age in children [81,82] and decreasing with age in adults [49,38,75,37,73,53,58,8,54]. A wide range of tongue strength values have been reported even in healthy populations, no doubt reflecting the influences of the age and sex of the population sampled. Values ranged from 49 to 73 kPa for males and 37 to 67 kPa for females. The analysis of younger (<60 years) compared to older adults indicated an average difference of 8 kPa for males and 9 kPa for females. There are likely to be differences between other age groups as well, but insufficient data exist at present to determine the magnitude of any differences. For future research studies, the age and sex effects on values mean that randomisation to groups should consider stratifying by age and sex.

For clinical practice, there is a need to develop sex-based normative data in a number of age bands, including children and adolescents. Also, a systematic investigation of tongue strength and endurance in adults and children with medical conditions is required as there are limited normative values for individuals with a medical condition.

Three studies conducted in healthy populations reported lower than typical tongue strength measures. Palmer et al. (2008) obtained much lower values during measurements obtained when intramuscular electrodes were inserted into specific muscles of the tongue. It is likely that the presence of the electrodes caused discomfort with muscle contractions altering their performance and reducing maximal strength performance. The second study by Vitorino (2010) examined tongue strength in Portuguese speakers with males (58.20 ±7.10 kPa) and females (57.10 ± 8.50 kPa) showing 11% lower tongue strength compared to English speakers. The inclusion of a small number (n = 10) of older Portuguese speakers may have contributed to the lower values as tongue strength has been shown to decrease in older people. The third study by Robbins et al. (1995) measured strength at different positions on the tongue (blade, dorsum, tip) in young (22 to 33y) and old (67 to 83y) healthy adults.

Despite the values being lower than those reported in many other studies the same trends were observed where older adults had lower tongue strength compared to younger adults.

3.8.2 Tongue endurance

Of the 16 evaluation studies in healthy participants, five measured tongue endurance, which was measured mostly in the anterior position. A wide range of values was observed, but there were no clear sex or age effects on tongue endurance. Two of the five included studies reported values lower than other studies included in this review. Vitorino (2010) reported mean tongue endurance as 16.20 ± 8.57 s. There is no clear explanation for these low values. Neel et al. (2010) reported values that were below the suggested normative range for males (37.85 ± 23.55 s) and for females (25.45 ± 14.11 s). Kays et al. (2010) reported endurance values measured in the anterior position for both males and females, but observed that lower endurance values were obtained from posterior measurements.

3.8.3 Hand strength

Few studies have reported hand strength measured by the IOPI. In general, males tend to have higher values than females, and younger adults higher values than older adults. Populations with a disorder also had lower hand strength values than healthy controls. There is a clear need for further studies to determine representative values for healthy sex-based age groups.

3.8.4 Hand endurance

Duration of hand endurance at 50% of maximum hand strength is not well established. Only one study investigated isometric hand endurance in only healthy individuals [38]. No significant sex or age effects were observed. Data from this study and the control group data in Table 5 indicate large variation in hand endurance values.

3.8.5 Studies in populations with a disorder

Most of the studies to date have been conducted in participants with PD, OPMD or head or neck cancer. Within each of these populations there are still too few data to gain a clear quantitative indication of the types of values that would be typical of these conditions. Most surprisingly, few studies have been conducted using the IOPI in stroke patients or many of the other neurological conditions. Thus, there is wide scope to establish IOPI values for tongue and hand strength and endurance in clinical populations.

3.8.6 Intervention studies

Five studies [87,17,18,35,88] used the IOPI as an evaluation tool in intervention research. Four of these studies [17,87,88,18] examined the effects of using the IOPI as a tongue-exercise training device, but no studies have used it as a training device for the hand. These studies clearly indicate that the IOPI can be an effective device for improving tongue strength, and possibly tongue endurance. There is now substantial scope to develop training protocols to address particular tongue strength or endurance deficits. The IOPI is also an effective tool to quantify the impact of tongue training interventions on tongue strength and endurance. There is also clearly potential to use the IOPI to track recovery after interventions or to provide better monitoring of loss of strength or endurance in progressive diseases.

The IOPI appears to be an effective tool to quantify the impact of tongue training interventions on tongue strength and endurance. Randomising participants to groups, including control groups, blinding the assessors, and performing and reporting sample size calculations could clearly improve the quality of reporting of these intervention studies. There is also room to improve the precision of measures by providing confidence intervals, or at a minimum, standard deviations. Also, the reporting of effects sizes would be beneficial to provide clear objective indications of the magnitude of any effects. Future studies should address these problems to prevent potential reporting bias.

3.9 Strengths and Limitations

There are several strengths to this review: the conduct and reporting of this review is aligned with the PRISMA statement for reporting of systematic reviews and meta-analyses; a comprehensive search strategy across multiple databases with no date restrictions; high agreement levels for quality assessments; and detailed data extraction to allow for comparisons between studies. However, the review also has some limitations. Unpublished literature was not located. This may have resulted in an over-representation of positive treatment effects (i.e., publication bias) in this review. Additionally, due to limited translation resources, only articles published in English were included. Therefore, it is possible that some studies addressing the use of the IOPI were not found. The studies investigating tongue and hand strength and endurance differed across many of the variables examined, including age groups; medical conditions; sex imbalance; study duration; group sizes; evaluation periods; exercise programs; IOPI bulb position; and training specificity. This inconsistency makes it difficult to determine the effect of these variables on outcomes and to compare effects across studies.

3.10 An application for clinical research and routine clinical practice

Based on the findings from this review, there is some evidence supporting the IOPI as an effective tool for research. The IOPI has primarily been used as an evaluation device, and it requires more investigation to determine its effectiveness as an intervention tool to improve strength or endurance for both adults and children with swallowing problems. There is enormous potential to improve patient outcomes in clinical practice by using a standardised assessment instrument such as the IOPI, which is relatively inexpensive and capable of providing objective measures of tongue strength and endurance rather than relying on the speech-language pathologist's clinical assessment, especially when multiple staff are making assessments. The IOPI has recently been approved by the Australian Therapeutic Goods

Administration for use in both research and clinical practice, which may increase the number of studies conducted outside the USA. There is a need to establish clearly relationships between tongue strength and endurance measures and swallowing function and performance in a range of populations. Also the reliability of these strength and endurance measures has not yet been reported.

3.11 Discussion

There is clear evidence indicating the effectiveness of the IOPI for the measurement of tongue and hand strength and endurance. This evidence is strongest for strength measurements, and is best established for measurements of tongue strength. There is a clear need to establish population specific representative values to gain maximum benefit from the use of these measures with this device.

Acknowledgements

The authors have declared that no competing interests exist and no author had any paid consultancy or any other conflict of interest with this article. This material was unfunded at the time of manuscript preparation.

Tables

Figures