

YouTube Exercise Videos for Lumbar Stenosis and Spondylolisthesis: a Scoping Review of their Quality, Reliability, and Validity

Hui Ling Li (Groningen, Pays-Bas), J.a. Bruinsma (Groningen, Pays-Bas), E.s. Van Ommen (Groningen, Pays-Bas), A.e.h. Broekema (Groningen, Pays-Bas), J.m.c. Van Dijk (Groningen, Pays-Bas), J.m.a. Kuijlen (Groningen, Pays-Bas), R. Soer (Groningen, Pays-Bas)

Introduction

Healthcare providers frequently advise patients diagnosed with lumbar stenosis or spondylolisthesis to improve their physical condition in preparation for surgical intervention. Online platforms, such as YouTube, are increasingly utilized as a self-management resource to access information and to perform exercises. Since low-quality videos might expose patients to misinformation or ineffective, sometimes even harmful exercises, this study aims to provide a comprehensive overview of the current quality, reliability, and validity of YouTube exercise videos for patients with a lumbar stenosis or spondylolisthesis.

Materials and Methods

A scoping review was conducted using a systematic YouTube search of four Dutch and four English terms. For each search term, the first 50 retrieved videos were reviewed. Exercise videos for lumbar stenosis or spondylolisthesis were included, while videos with explicit contraindications, inadequate audio/visual quality, and unrelated content, including postoperative and traumatic cases, were excluded. *Quality* was evaluated using the Global Quality Scale (GQS; range 1-5) and modified DISCERN (mDISCERN; range 1-5); *Inter-rater reliability* was analysed using Cohen's kappa coefficient (κ); *Video reliability* was assessed with the Journal of the American Medical Association benchmark criteria (JAMA, range 0-4); *Validity* was determined using the content and disclosed references of each video.

Results

Of 400 retrieved videos, 99 videos were included in the analysis (53 stenosis, 46 spondylolisthesis) after removing duplicates ($n=64$) and excluding videos based on title ($n=203$) or predefined criteria ($n=34$). Physiotherapists were the leading group of uploaders, accounting for 42 (79.2%) videos on stenosis and 24 (52.2%) videos on spondylolisthesis. *Quality*: The median GQS and mDISCERN score was 3 (moderate quality) for both conditions. *Reliability*: High content reliability (JAMA ≥ 3) was observed in 34 (64.2%) videos on stenosis and 24 (52.2%) videos on spondylolisthesis. *Inter-rater reliability* was low for GQS ($\kappa = 0.48$ and 0.27 , respectively), moderate for mDISCERN ($\kappa = 0.63$ and 0.54 , respectively), and high for JAMA ($\kappa = 0.86$ and 0.81 , respectively). *Validity*: Validity was limited; only 3 (3.0%) videos cited sources to support the recommended exercises.

Conclusion

The majority of YouTube videos on exercises for lumbar stenosis and spondylolisthesis were of moderate quality, variable reliability, and limited validity. Although many videos aligned with existing literature, notable discrepancies were observed, including the limited representation of certain effective, evidence-based exercises and the inclusion of potentially harmful exercises. Evidence-based videos of higher quality and reliability are needed to support preoperative self-management in patients with lumbar stenosis and/or spondylolisthesis.

Table 1. Distribution of quality, reliability, and validity

			Stenosis (n=53)	Spondylolisthesis (n=46)	p-value	
Quality						
GQS, n (%)	Very low		3 (5.7%)	3 (6.5%)	0.22	
	Low		14 (26.4%)	4 (8.7%)		
	Moderate		23 (43.3%)	24 (52.2%)		
	High		10 (18.9%)	13 (28.3%)		
	Excellent		3 (5.7%)	2 (4.3%)		
	Median (IQR)		3 (2 - 3.5)	3 (3 - 4)		
mDISCERN, n (%)	Very low		5 (9.4%)	3 (6.5%)	0.01*	
	Low		19 (35.8%)	9 (19.6%)		
	Moderate		22 (41.5%)	17 (37.0%)		
	High		4 (7.5%)	16 (34.8%)		
	Very high		3 (5.7%)	1 (2.2%)		
	Median (IQR)		3 (2 - 3)	3 (2 - 4)		
	Item					
		Aims clear	38 (82.6%)	38 (82.6%)		
		Reliable sources	42 (91.3%)	42 (91.3%)		
		Balanced and unbiased	2 (4.3%)	2 (4.3%)		
		Additional sources	23 (50.0%)	23 (50.0%)		
		Areas of uncertainty	35 (76.1%)	35 (76.1%)		
Reliability						
Inter-rater reliability, κ		GQS	0.48 (low)	0.27 (low)		
		mDISCERN	0.63 (moderate)	0.54 (moderate)		
JAMA, n (%)	Low	JAMA	0.86 (high)	0.81 (high)	0.52	
		1	3 (6.5%)	3 (6.5%)		
		2	19 (41.3%)	19 (41.3%)		
	High	3	22 (47.8%)	22 (47.8%)		
		4	2 (4.3%)	2 (4.3%)		
	Median (IQR)	3 (2 - 3)	3 (2 - 3)			
	Criteria	Authorship	36 (78.3%)	36 (78.3%)		
		Attribution	2 (4.3%)	2 (4.3%)		
		Disclosure	31 (67.4%)	31 (67.4%)		
		Currency	46 (100.0%)	46 (100.0%)		
Validity						
Content ^{a,b}	Strength		34 (64.2%)	42 (91.3%)	<0.01*	
	Flexibility		47 (88.7%)	16 (34.8%)	<0.01*	
	Coordination	Core stability	34 (64.2%)	34 (73.9%)	0.30	
		Balance	0 (0.0%)	2 (4.3%)	0.21	
	Lumbar/Spinal	Stabilization	34 (64.2%)	33 (71.7%)	0.42	
		Flexion	17 (32.1%)	1 (2.2%)	<0.01*	
		Extension	0 (0.0%)	1 (2.2%)	0.47	
	Pain-relieving	19 (35.8%)	6 (13.0%)	0.01*		
	Adjustments	7 (13.2%)	10 (21.7%)	0.26		
	References	Available	Exercise	2 (3.8%)	1 (2.2%)	
			Other	3 (5.7%)	1 (2.2%)	
		Not available		48 (90.6%)	44 (95.7%)	

Abbreviations: GQS, Global Quality Scale; mDISCERN, modified DISCERN; JAMA, Journal of the American Medical Association

Statistical tests: Pearson Chi-Square test for GQS, mDISCERN, JAMA, and content (strength, flexibility, core stability, stabilization, flexion, pain-relieving, and adjustments). Fisher's exact for content (balance and extension).

(*) indicates a significant difference of p<0.05

(a) Some videos were assigned to more than one category

(b) There were no videos presenting endurance and speed exercises.