



Supplement 1. Changes in patients self-reported outcomes score in subgroup of patients with complete core outcomes measure index (COMI) data at all four timepoints (n=247). Median and interquartile range of overall COMI score at each timepoint shown. **(A)** Change in overall COMI score over time. COMI scores were lower at 3 months, 1 year, and 2 years post-surgery compared to pre-treatment ($p < 0.001$, mixed-effects model). The largest decrease in COMI score was seen in the first 3 months post-surgery. p -value shown represents difference from pre-surgery COMI score and calculated using a repeated measures mixed-effects model with imaging time-point as a fixed effects variable. *Post hoc* analysis of pairwise comparisons between different timepoints was performed using the Bonferroni method. **(B)** Change in overall COMI score stratified by lumbar fusion approach. p -value shown represents differences in COMI score between surgical approaches at each timepoint. p -value calculated using Kruskal-Wallis test. PLIF-BMS, posterior lumbar interbody fusion with midline posterior approach/bilateral muscle strip; PLIF-Wiltse, PLIF with bilateral Wiltse/muscle splitting approach; MI-TLIF, minimally invasive transforaminal lumbar interbody fusion; Open TLIF, open transforaminal lumbar interbody fusion; PLF-only, pedicle screw/posterolateral fusion only; Preop, preoperative. * $p \leq 0.05$. *** $p \leq 0.001$.

Supplement 2. Patient self-reported outcomes stratified by lumbar fusion approach

Procedure	All (N=247)	PLIF bilateral muscle strip (N=116)	PLIF Wiltse approach (N=65)	MI-TLIF (N=43)	Open TLIF (N=12)	Posterolateral fusion only (N=11)	p -value
Preop COMI score	8.10 (7.0–9.5)	8.30 (6.9–9.5)	8.35 (7.1–9.0)	8.10 (7.0–9.75)	7.95 (6.3–9.4)	7.60 (7.0–9.3)	0.95
3-Month postop COMI score	5.30 (3.2–7.0)	5.40 (3.0–7.1)	4.90 (3.4–6.70)	5.40 (3.2–6.8)	4.42 (3.4–5.9)	5.65 (5.3–7.6)	0.61
1-Year postop COMI score	4.15 (2.1–6.3)	4.70 (2.1–6.9)	3.00 (1.3–5.2)	5.20 (2.1–7.1)	5.30 (2.6–6.6)	6.25 (3.3–7.6)	0.03
2-Year postop COMI score	4.00 (2.0–6.2)	4.45 (2.0–6.6)	3.10 (1.6–5.0)	5.00 (2.3–8.1)	4.80 (2.0–6.5)	5.00 (4.0–7.9)	0.06

Values are presented as median (interquartile range). Subgroup analysis of patients with complete COMI data at all four timepoints (n=247). For differences between lumbar fusion approaches, p -value calculated using Kruskal-Wallis test. Significant results ($p < 0.05$) shown in bold.

PLIF, posterior lumbar interbody fusion; MI-TLIF, minimally invasive transforaminal lumbar interbody fusion; Open TLIF, open transforaminal lumbar interbody fusion.