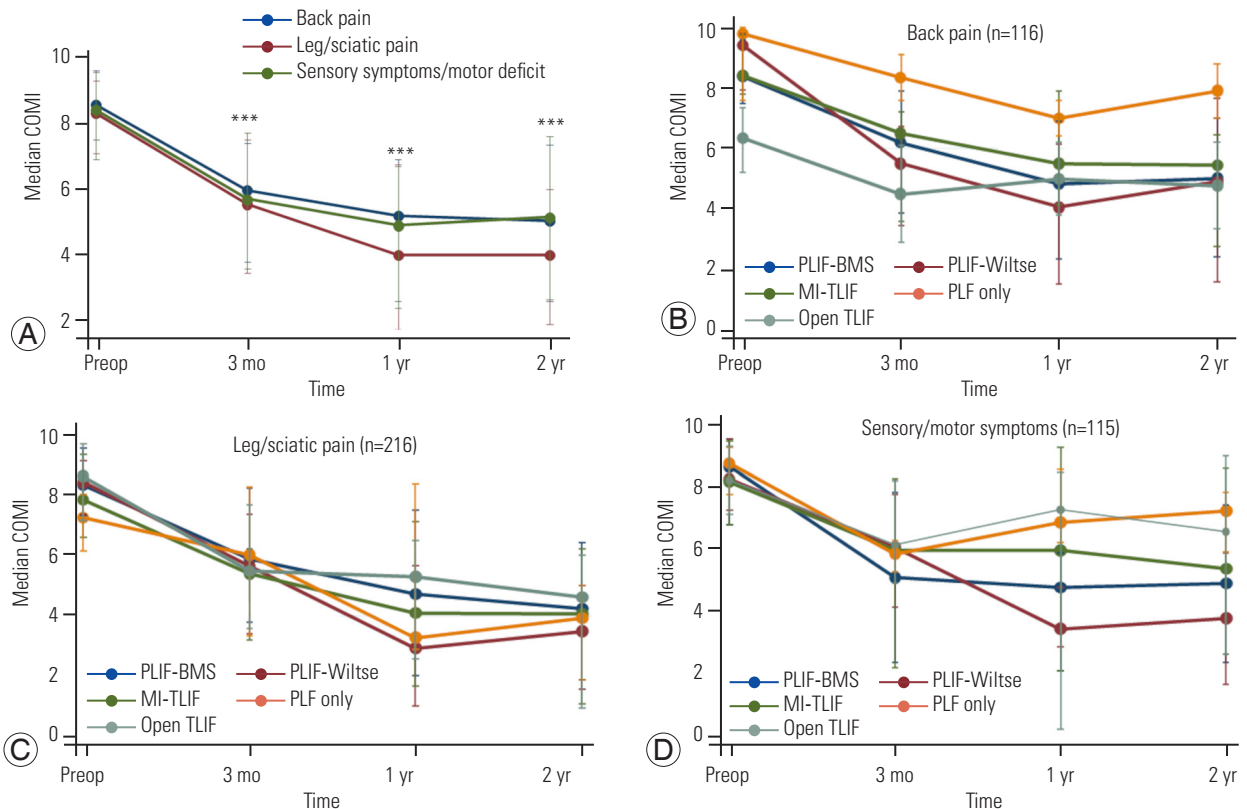


Supplement 3. Changes in patients self-reported outcomes score in subgroup of patients undergoing surgery for degenerative spondylolisthesis only ($n=226$). Median and interquartile range of overall core outcomes measure index (COMI) score at each timepoint shown. (A) Change in overall COMI score over 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. Only patients undergoing bilateral cage insertion included in PLIF procedure groups. 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 4. Changes in patients self-reported outcomes stratified by primary (predominant). Symptom indication median and interquartile range of overall core outcomes measure index (COMI) score at each timepoint shown. (A) Change in overall COMI score over time. Curves shown for each predominant symptom indication (back pain, leg/sciatic pain, or lower limb sensory symptoms/motor deficit). 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–D) Change in overall COMI score stratified by lumbar fusion approach. Graphs shown for each predominant symptom indication. 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. *** $p\leq 0.001$.