Partial intervertebral fusion secures successful outcomes after thoracoscopic anterior scoliosis correction; a low dose computed tomography study

Maree T Izatt (BPhty), Alan Carstens (FRACS), Clayton J Adam (PhD),
Robert D Labrom (MSc, FRACS), Geoffrey N Askin (FRACS)

Paediatric Spine Research Group, Institute of Health and Biomedical Innovation, Queensland University of
Technology and Mater Health Services, Brisbane, Australia

Name and Address for Correspondence

Maree Izatt
Paediatric Spine Research Group
Level 5, Centre for Children’s Health Research
62 Graham Street, South Brisbane, Queensland, 4101, Australia
Phone: +61 7 3069 7324 Email: m.izatt@qut.edu.au
Abstract

Study Design
Retrospective review of prospectively collected data

Objectives
To analyze intervertebral (IV) fusion after thoracoscopic anterior spinal fusion (TASF) and explore the relationship between fusion scores and key clinical variables.

Summary of background information
TASF provides comparable correction with some advantages over posterior approaches but reported mechanical complications and their relationship to non-union and graft material is unclear. Similarly, the optimal combination of graft type and implant stiffness for effecting successful radiologic union remains undetermined.

Methods
A subset of patients from a large single centre series who had TASF for progressive scoliosis underwent low dose CT scans two years after surgery. The IV fusion mass in the disc space was assessed using the four point Sucato scale, where 1 point indicates <50% and 4 points indicates 100% bony fusion of the disc space. The effect of rod diameter, rod material, graft type, fusion level, and mechanical complications on fusion scores were assessed.

Results
43 patients with right thoracic major curves (mean age 14.9 years) participated in the study. Mean fusion scores for patient subgroups ranged from 1.0 (IV levels with rod fractures) to 2.2 (4.5mm rod with allograft), with scores tending to decrease with increasing rod size and stiffness. Graft type (autograft versus allograft) did not affect fusion scores. Fusion scores were highest in the middle levels of the rod construct (mean 2.52), dropping off by 20-30% toward the upper and lower
extremities of the rod. IV levels where a rod fractured had lower overall mean fusion scores compared to levels without a fracture. Mean total SRS questionnaire scores were 98.9 from a possible total of 120, indicating a good level of patient satisfaction.

Conclusions

Results suggest that 100% radiologic fusion of the entire disc space is not necessary for successful clinical outcomes following thoracoscopic anterior selective thoracic fusion.

Keywords

thoracoscopic scoliosis surgery; intervertebral fusion; computed tomography; rod diameter; fusion grading; interbody fusion; anterior spinal surgery; adolescent idiopathic scoliosis; complications.
**Introduction**

Thoracoscopic anterior spinal fusion (TASF) is an effective surgical technique in selected cases for the treatment of progressive adolescent idiopathic scoliosis (AIS) [1-7]. Potential benefits of this technique include fewer fused levels, restoration of the sagittal profile, reduced pain and chest wall morbidity, shorter hospital stay, faster recovery of lung function, improved cosmesis and a lower infection rate, blood loss, implant density and incidence of neurological complications [8-15].

However, mechanical complications such as screw pullout or rod fracture have been well documented in the range of 9 -16% for this minimally invasive single rod technique [2, 4, 6, 7, 11, 16-18]. The development of a non-union following instrumented scoliosis correction may predispose to curve progression, pseudarthrosis and subsequent implant failure [4, 19, 20]. Some curve progression has been identified in the first two years after surgery [4, 6, 18], but the relationship of this small progression to instrumentation complications and the possible causative role of non-union of the arthrodesis remain unclear. Furthermore, the relationship between radiologic union and clinical outcome is not clear, as apparent non-union of a spinal fusion, rod breakage or top screw pullout does not always correspond to a poor clinical result, the need for revision surgery or patient dissatisfaction [7, 21, 22]. Thirdly, although prior studies [17, 23] have suggested a link between implanted rod diameter and the maintenance of deformity correction and/or the incidence of non-union and rod fracture in TASF, the optimal combination of graft type and implant stiffness for effecting a successful radiologic union remains undetermined.

The aim of this study was to perform the first quantitative thin slice computed tomography (CT) analysis of two-year postoperative intervertebral fusion in a group of patients who had received TASF. The results were analyzed in order to explore the relationship between fusion scores and; rod
diameter and stiffness, graft material, fusion level, and the occurrence of postoperative mechanical complications.

**Materials and Method**

**Study Cohort.**

Between 2005 and 2008, a subset of 43 patients from a large single centre consecutive series of 210 patients who have undergone thoracoscopic anterior single rod fusion, consented to participate. The CT scans were gathered prospectively after seeking and obtaining ethics committee approval from our institution to perform low dose CT scans at a minimum of 24 months after surgery on a subset of patients who had TASF. All patients who reached a minimum 24 months follow-up after surgery, during the ethical approval period, were invited to participate. Patients who had reached their 24 month follow-up prior to the ethics approval being granted, and had a known mechanical complication (top screw pullout or rod fracture), were identified from the clinical TASF database and also invited to participate. The study group was therefore not a representative sample of the whole series, with deliberate targeting of patients with mechanical complications. However, this study design allowed assessment of the fusion scores achieved at each vertebral level in those who had mechanical complications compared to participants without mechanical complications. Of the 40 TASF patients who reached 24 months follow-up during the ethics approval period, 32 (80.0%) agreed to participate. All patients who had reached 24 months follow up prior to the ethics approval and had a complication (14 patients) were contacted, and 11 (76.9%) agreed to participate.

**Surgical Technique.**

The surgeries were performed by the senior authors (GNA and RDL) at the Mater Children’s Hospital in Brisbane, Australia. The thoracoscopic surgical procedure is based on the technique first described
by Picetti *et al* [1] and has been reported previously [11, 24, 25]. The cleared disc spaces of the vertebral levels to be instrumented were packed with either autograft (patients 1-13) or allograft (patients 14-43). The change from autograft (which was used for the first 45 cases of the full patient series at our centre) to irradiated mulched femoral allograft supplied through the Queensland Bone Bank [6] was due to donor site pain and the difficulty accessing the posterior iliac crests from the side lying position (iliac crest) and the inadequate volume of bone available from ribs in this patient cohort. Twenty-three patients in the current study had a single 4.5mm diameter pure titanium rod implanted and the following 20 more recent patients had a 5.5mm diameter rod (7 titanium alloy initially followed by 13 pure titanium, which is now the standard practice). The changes in rod diameter and then rod materials occurred in the full TASF cohort in an effort to reduce the rate of rod breakage which occurred in the earlier part of the larger thoracoscopic series [6]. The levels chosen to be instrumented were selected to include the end vertebrae of the major scoliotic curve. Radiographs and photographs of a typical patient before and at two years after TASF are shown in Figure 1.

**Low dose CT Evaluation**

The size, location and quality of the fusion mass and implant integrity was evaluated at a minimum of two years following surgery via a single thoracolumbar CT scan using a low-dose scanning protocol. Three different CT scanners were used during the three year data collection period of the study; (i) a 64-slice Philips Brilliance (Philips Healthcare, Andover, USA), (ii) a 64 slice GE Lightspeed Plus (GE Healthcare, Chalfont St. Giles, UK) and (iii) a 64 slice GE Lightspeed VCT (GE Healthcare, Chalfont St. Giles, UK). The scan coverage in each case was from C7 to S1. Dose reports were commissioned for all three scanners, and the highest estimated radiation dose of 3.0mSv occurred with the GE Lightspeed VCT scanner, with uncertainties due to the dose model in the order of ±20% [26]. Estimated doses for the other 64 slice scanners were substantially lower, in the order of 2.0mSv.
By comparison, the combined dose for a postero-anterior and lateral standing radiograph is approximately 1.0mSv and the annual background radiation in Queensland, Australia is approximately 2.0mSv.

Reformatted sagittal plane images were produced from the transverse CT slices using the ImageJ image processing software (v. 1.42q, National Institutes of Health, Bethesda, USA) which allows the entire scoliotic spine to be visible on a single sagittal image. Due to the lateral curvature of the scoliotic spine, the whole spine is not generally visible in a single planar sagittal image. Three reformatted sagittal plane images were produced for each participant at three locations in relation to the implanted rod\(^1\). Typical reformatted sagittal images using this technique are shown in Figure 2, (a) close to the rod (right parasagittal), (b) at the midpoint of the vertebrae (mid-sagittal) and (c) furthest from the rod (left parasagittal), for a single patient. The window and level of the resulting greyscale images was adjusted to display a range of Hounsfield Units between approximately 100 to 1000, in order to show mineralized tissue densities between that of vertebral trabecular bone \cite{27} and cortical bone \cite{28} for the subsequent fusion grading. The reformatted sagittal plane images for each participant were de-identified and supplied to three observers electronically. The independent observers (two spinal fellows and an experienced research assistant) were instructed to grade the images on two separate occasions at least two weeks apart. The observers were blinded to participant identity, their previous scores and the image order was randomized for each grading occasion. The observers were instructed to indicate if any intervertebral levels were unable to be graded due to image quality. The intervertebral levels fused as part of the surgical correction were numbered from 1, (cephalad to caudad) regardless of the vertebral levels chosen by the surgeon, to allow analysis of the fusion quality and mass along the length of the instrumentation between participants.

\(^1\) Since all patients in the study had right thoracic major scoliotic curves, the single anterior rod was always placed on the right side of the vertebral column.
The Sucato method [29] was used to grade the fusion on the left para-sagittal, right para-sagittal and mid-sagittal CT reconstructions for each intervertebral level. The Sucato fusion grading method uses a 4-point grading scale, with 0 points indicating no fusion mass; 1 point indicating fusion of <50% of the area of the disc space; 2 points indicating fusion of between 50% and 75% of the area of the disc space; 3 points indicating fusion of >75% of the area of the disc space; and 4 points indicating complete fusion across the entire disc space (Figure 3). The percentage of disc space fusion is calculated by dividing the osseous fusion area by the total discectomy area enclosed within the proximal and distal end plates and the posterior and anterior vertebral body margins. A score of 3 or 4 points was recommended by the Sucato method to represent a solid fusion. In the current study, once all instrumented disc spaces in all patients had been graded, the resulting fusion scores were analyzed to look for relationships between the fusion scores and the following four possible influencing variables; (i) rod diameter/stiffness, (ii) graft material, (iii) fusion level, and (iv) occurrence of post-operative mechanical complications.

Patient satisfaction and quality of life after TASF was evaluated using the SRS-24 questionnaire at 24 months after surgery which had been collected prospectively and results stored in the clinical TASF database.

**Statistical analysis.**

The mean and standard deviation of fusion scores for all instrumented vertebral levels were calculated from the fusion scores of all three observers. All the fusion grading scores were then analysed for the four subgroups listed above. The particular IV levels where a rod fracture or screw pullout occurred were excluded from all sub-group analyses that were unrelated to mechanical complications. Inter-
Intervertebral fusion following anterior scoliosis surgery

and intra-observer variability for fusion grading was assessed using the approach described by Bland and Altman [30, 31]. Intra-observer variability was assessed by analyzing the absolute difference between successive fusion scores of the same intervertebral level (α) calculated as:

$$\Delta \alpha = |\alpha_n - \alpha_m|$$

where n and m are successive fusion scores made by the same observer. The 95% confidence intervals for intra-observer variability were calculated as $1.96 \times SD_{\text{intra}}$ [30, 31] where $SD_{\text{intra}}$ is the standard deviation of the intra-observer differences $\Delta \alpha$.

The inter-observer variability (standard deviation of the difference between fusion gradings by different observers) was calculated as $\sqrt{2} \times SD_{\text{inter}}$ for a single fusion score per observer, where $SD_{\text{inter}}$ is the standard deviation of the inter-observer differences. The 95% confidence intervals for inter-observer variability were calculated as $1.96 \times SD_{\text{inter}}$ ($t$-distribution with 42 dof).

Independent samples Kruskal-Wallis tests were performed using SPSS (Version 22.0 Armonk, NY: IBM Corp) to test for differences in overall fusion score versus patient skeletal maturity at the time of surgery (Group 1= 11 patients Risser 0 or 1, Group 2= 13 patients Risser 2 or 3, Group 3= 19 patients Risser 4 or 5), and for fusion score versus two year post-operative major Cobb angle (Group 1= 30 patients Cobb 10-24 degrees, Group 2= 8 patients Cobb 25-34 degrees, Group 3= 5 patients Cobb 35 degrees or more). Tests were performed both with and without inclusion of IV levels demonstrating instrument related complications in the cohort.
**Results**

A total of 43 patients (39 females, 4 males) participated in the study. Low dose CT data for all patients was available at minimum 22 months (mean 2.7, SD 1.3 years, range 1.8-6.2) after their surgical corrections which had been performed between 2000 and 2006. The study cohort has been followed up for mean 10.7 years (range 8.2 – 14.7). The mean age at surgery was 14.9, SD 3.1 years (range 9.9-27.8), with 3 patients aged over 18 years demonstrating sufficient major and compensatory curve flexibility to be considered suitable for this selective anterior thoracic fusion procedure. Of the study group, all 43 had right thoracic major curves classified as Lenke class 1A (n=24), 1B (n=13) or 1C (n=6). Main thoracic Cobb angle before surgery was mean 51.0 degrees (range 40-66) which was corrected to mean 22.2 degrees (range 10-42) after surgery.

The mean number of levels instrumented was 6.9, SD 0.7 (range 5-8), and mean operative time was 270.4, SD 63.0 minutes (range 185-430). The proximal extent of the instrumentation was to T5 (n=13), T6 (n=25) or T7 (n=5). The distal instrumented levels were to T10 in 2 cases, T11 (n=9), T12 (n=30) and L1 in 2 cases. Within the study cohort 9 (20.9%) patients were braced for 12 weeks, 5 were braced for 8 weeks (11.6%), and 10 (23.3%) were braced for 6 weeks after surgery. The most recent 19 (44.2%) patients were not braced after surgery, in line with current practice. These changes in bracing protocol represent a gradual decline in bracing time, from the initial patients in the larger thoracoscopic series, to the current practice of no bracing after surgery, based on increasing experience and confidence of the surgical team in the procedure [11].

In the current study cohort, including the deliberately targeted complication cases, there were 11 cases of rod fracture and 6 unrelated cases with screw-related complications (Figure 4). It is important to note that these complication numbers are not indicative of the prevalence of mechanical complications
in the larger fully inclusive TASF series. Of the 32 patients who reached 24 month follow up during the study period, 1 rod fracture was found and 5 of the 6 screw related complications (5 partial screw pullouts and one screw plow). The additional 10 rod fracture cases were from the deliberately targeted patients invited from the full TASF series (currently 210 cases) who had already reached 24 months follow up, prior to the commencement of the study. It is important to note, that although the fusion analysis was performed on CT taken at mean 2.7 years (range 1.8 – 6.2) after surgery, the study cohort has now reached a mean follow up of 10.7 years (range 8.2 – 14.7) to ensure that later instrument complications were noted and included in the sub group analyses. The rod fracture rate in the fully inclusive TASF patient series at our centre currently stands at 10%, with a screw related complication rate of 5.2% with the follow-up range 2.0 - 14.7 years. Previous papers published on larger cohorts from our single centre series [6, 7] (106 and 100 patients respectively) report radiographic and clinical outcomes and complication rates for larger and fully inclusive cohorts to that of the current study cohort. These studies report the incidence of rod breakage at 11.3 – 13.0% and screw pullout/plow 7.0 - 8.5%. All the rod fractures occurred after the 12 month review and 6 of the 11 were present at the scheduled 24 month review, two at 36 months, and the final 3 occurred between 5 and 7 years postop (despite the patients being asymptomatic). For 10 out of the 11 rod fracture cases in the study cohort, the rod break was found near the two apical IV segments and in no cases did a rod break closer than three levels to either end of the construct. All the screw related complications occurred in the early postoperative period and were noted either prior to leaving hospital or at the scheduled 6-8 week review appointment. No patient had both a rod fracture and a screw related complication in the current study (nor in the full TASF series of 210 to date).
Assessment of the 43 CT scans by three observers on two occasions resulted in a total of 4023 fusion scores measured in 256 disc spaces. The Sucato scores for all the subgroups analysed are summarized in Table 1.

(i) Effect of rod diameter/stiffness on fusion scores

Rod diameter had a strong effect on fusion score, with a mean score of 2.18, SD 1.33 for the 4.5mm Titanium rod, decreasing significantly to 1.43, SD 1.09 (P=0.016, unpaired t-test) for the 5.5mm pure titanium rod, and to 1.10, SD 0.92 for the (stiffer) 5.5mm titanium alloy rod, when the scores for all three sagittal reconstructed planes and all fusion levels in a particular patient were combined. The difference in fusion score between 4.5mm pure titanium rod and 5.5mm titanium alloy rod was also significant (P=0.003), however the fusion scores for the pure and alloy 5.5mm rods were not statistically different (P=0.19). Note that all of these rod comparisons were for allograft participants.

When the three sagittal reconstructions were analysed independently the trend towards lower fusion scores with increasing rod size and stiffness was consistently repeated in all three reconstruction planes, see Figure 5.

(ii) Effect of graft type on fusion scores

Mean fusion scores for autograft and allograft subgroups when a 4.5mm rod diameter was used were similar at 2.15 SD 1.34 vs 2.22 SD 1.32 respectively (P=0.96). See Figure 6.

(iii) Effect of fusion level on fusion scores

Mean fusion scores were highest in the middle intervertebral levels (IV levels 2,3,4 had mean score of 2.52 closest to the rod) of the implant construct for all three sagittal reconstructed planes, dropping off by 20-30% toward the upper and lower extremities of the rod. When combining left, right and mid-
sagittal fusion scores for a particular vertebral level, the difference in fusion scores between adjacent levels was only statistically significant between the 4th to 5th levels in the construct (P=0.001). Figure 7 shows a graph of the left, right and mid-sagittal fusion scores versus fusion level.

*(iv) Lateral variation in fusion scores*

The overall mean fusion scores closest to the rod, i.e., the right para-sagittal reconstruction (2.28, SD 1.18) were significantly higher than on the contralateral side of the disc space (left para-sagittal reconstruction and furthest from the rod) 1.39, SD 1.33 (P<1×10^-6 for both left para-sagittal vs mid-sagittal, and mid-sagittal vs right-parasagittal, paired t-test). This pattern of increasing fusion scores closest to the rod is evident in all subgroup analyses (see Table 1).

*(v) Effect of post-operative mechanical complications on fusion scores*

Intervertebral levels where a rod fracture occurred (n=11) had statistically significant lower overall mean CT fusion scores compared to those levels without a fracture (1.00, SD 0.93 versus 1.90, SD 1.31) (P=0.001, Mann-Whitney U test). When analysed in each of the three sagittal planes, the difference became more marked as the reconstruction became closer to the rod, suggesting that the important determinant of rod fracture was the fusion score close to the rod, rather than that on the contralateral side (Figure 8).

Intervertebral levels immediately below where a top screw pullout occurred (i.e. the first fused level, n=6) had lower overall mean CT fusion scores when compared to those patients without complications (1.35, SD 0.94 versus 1.95, SD 1.32), although this difference was not statistically significant (P=0.083, Mann-Whitney U test). Despite these identified mechanical complications there was only one patient who required a revision procedure seven years after surgery. This patient
reported discomfort and some curve progression was noted on the coronal radiograph. The broken rod was replaced using the existing vertebral body screws and the same access portals. The pseudarthrosis accompanying the rod fracture was cleared and packed with Bone Morphogenetic Protein (BMP-2, Infuse, Medtronic, USA) and local bone, which achieved excellent correction which has been maintained in the six years since this revision procedure. Other participants have continued to be monitored in the longer term, with the most recent now eight years after surgery, and all curves have remained stable.

(vi) Inter and Intra-observer variability in fusion scores

Inter and Intra-observer variability for fusion grading at each intervertebral level is given in Table 2. The mean difference between successive fusion gradings by the same observer was not significantly different from zero, and suggests that no order bias existed between the first and second gradings of an intervertebral level.

(vii) Effect of skeletal maturity and Cobb angle

There were no significant differences in fusion score versus patient skeletal maturity for either the entire patient cohort (P=0.15), or for the subgroup of patients with no complications (P=0.192). There were no significant differences in fusion score versus post-operative major Cobb angle for either the entire patient cohort (P=0.128) or for the subgroup of patients with no complications (P=0.614)

(viii) Patient Satisfaction

SRS questionnaires were available for 42 (96%) patients. The mean total SRS score for all patients was 98.9, SD 7.3 (range 82 – 110) from a possible total of 120 points, indicating a favourable clinical outcome for the procedure. The patients with no mechanical complications (n=25), rod fracture
(n=11) or top screw pullout (n=6) scored mean SRS scores of 98.0, 101.5 and 98.0 points respectively. Mean scores for all domains are shown in Table 3.

**Discussion**

The objectives of correction surgery in progressive AIS are to permanently halt progression, correct the deformity in three dimensions, and ultimately to improve trunk appearance. To our knowledge, this is the first detailed investigation of CT generated fusion scores after selective anterior thoracic fusion. The analysis of the fusion mass location and size in the cleared intervertebral disc space provides new information to consider when choosing the optimal implant type and interbody graft material to achieve clinically effective fusion in single rod anterior scoliosis correction surgery. Increasing rod stiffness for the construct by increasing rod diameter and alloy properties had the effect of reducing fusion scores in our patient group (Figure 5). These results suggest that a larger diameter rod may require a smaller fusion mass to achieve a stable intervertebral segment. Importantly, the current study data indicates that the fusion mass does not necessarily need to be large, nor fill the entire intervertebral disc space to effect satisfactory fusion in single rod anterior fusion procedures for idiopathic scoliosis patients.

The graft type sub-analysis (Figure 6) found that graft type does not affect the resulting fusion score. The graft types analysed in this study involve two very different methods with respect to sourcing and preparation of the bone graft material prior to placement. Rib or iliac crest autograft harvested at the time of surgery (in the earlier cases of the larger series and the study cohort) is limited in supply and in the case of iliac crest graft, is very difficult to obtain with the patient positioned in side lying for the thoracoscopic anterior approach. Mulched femoral allograft is available in copious amounts and also avoids the problematic postoperative donor site pain of autograft harvesting. However, the allograft
does not contain osteo-inductive agents which have been shown to enhance fusion potential [32-34]. The fact that no difference in fusion score was found in this study suggests that these differences do not result in any demonstrable change in fusion scores between the graft types. A recent study by Lee et al [18] also concluded that the graft type used in thoracoscopic anterior correction did not affect the incidence of pseudarthrosis and subsequent implant failure.

The fusion scores consistently peaked in the middle intervertebral disc spaces (Figure 7), where surgical intervertebral compression forces are highest surrounding the apex of the scoliotic deformity. Relative to these apical levels, fusion scores were reduced by 20-30% in the highest and lowest two levels where the lateral curves tend to flatten out. A recent finite element modelling study by Little et al [35] found that the majority of the deformity correction in TASF occurred in the intervertebral disc spaces at or near the apex of the deformity. Further work is required to investigate whether there is a relationship between surgical corrective forces and the resulting fusion scores for this patient group.

For all levels, the mean fusion scores were higher on the side of the disc space near the rod, than on the contralateral side (Figure 7). This finding seems intuitive given the nature of the surgical technique and minimally invasive approach to the anterior spinal column. The thoracoscopic approach to the spine does not allow a full soft tissue intervertebral disc clearance at the levels to be instrumented. Due to access, clearance is maximal at the apex of the curve and at the side of thoracoscopic portal placement or the convexity of the curve. A variable quantity of disc material will remain at some levels which makes it difficult to obtain symmetrical (across disc space and over the central and proximal/distal disc spaces) placement of bone graft. Additionally, the convex side of the spine which is the instrumented side is preferentially loaded in compression as part of the surgical technique which may also contribute to the formation of bone on this convex side. It may also be that
the degree of micro-motion during post-operative activity is conducive to bone formation on the instrumented side, but is too high for bone formation on the contralateral side of the cleared disc space, leading to fibrous tissue formation instead. However, from a clinical perspective (post-operative maintenance of correction as well as SRS questionnaire scores), satisfactory fusion is achieved using the TASF technique. Therefore we conclude that partial interbody fusion (fusion grades of around 2 on the Sucato scale representing 50-75% filling of the disc space) are sufficient to secure successful outcomes after TASF.

The Sucato fusion grading method was first described in 2004 in order to undertake anterior interbody fusion assessment in a thoracoscopically instrumented porcine model [29]. It is a score based on the percentage of bony fusion across the entire disc area using CT scan imaging and to date has only been used in animal models. The Sucato scale was used in the present study as the only validated fusion grading assessment method found in the literature for the analysis of anterior interbody fusion following thoracoscopic scoliosis correction. As discussed above, fusion scores of around 2 on the Sucato scale secure satisfactory clinical outcomes, yet a score of ‘only’ 2/4 could imply inadequate fusion. However due to the nature of the thoracoscopic technique it is extremely difficult to achieve a full discectomy and clearance at each level and therefore a fusion mass which fills the entire intervertebral space [3, 36]. Yet this study has shown that the ‘spot weld’ region of high bone density usually found on the side nearest the rod (convex side) in the thoracoscopic technique may be sufficient for a satisfactory fusion result, thereby avoiding rod breakage and subsequent progression of the instrumented curve. For these reasons, we suggest that the Sucato method used in the current study is not ideal to judge the fusion achieved with minimally invasive anterior approaches. An alternative grading system for fusion assessment could be formulated to suit this approach. In particular, a future fusion grading system could take advantage of the capability of CT grayscale values to provide
information on bone density, and therefore capture the distribution of bone quality in the disc space rather than the volume of fusion mass throughout the entire disc space. Further topographical analysis of the fusion mass in three planes (frontal and transverse as well as the sagittal plane results presented here) would also aid the ability to grade fusion compared to uni-planar assessment of the reconstructed sagittal images as described by Sucato.

Postoperative care protocols involving bracing patients after fusion surgery vary amongst surgeons and institutions but the literature suggests that the majority of surgeons continue to brace patients for three months after TASF surgery [2, 4, 9, 18]. The most recent 135 cases in the larger TASF series of 210 cases at our Centre have not been braced at all after surgery, due to increasing confidence in the surgical technique, with no resulting increase in mechanical complications or incidence of pseudarthrosis to date. The bracing protocol for the patients in the current study was coincidentally changed at the same time as the rod diameter. This makes it impossible to comment on the effect of bracing in isolation in effecting a clinically satisfactory fusion, but based on our subsequent anecdotal experience since bracing has been discontinued; we suggest that postoperative bracing does not influence fusion grade or clinical success. Non-parametric tests showed that fusion scores were also not significantly affected by patient skeletal maturity at the time of surgery, nor by the size of the deformity after surgical correction (i.e. the 2 year post-operative Cobb angle).

The mostly apical levels where a rod fractured had a lower overall mean CT fusion score than those IV levels where the rod was intact (Figure 8). This supports the commonly held theory of rod fracture occurring as a fatigue failure due to inadequate fusion at a particular IV level. Of the 11 levels where the rod fractured, 8 occurred in the 4.5mm rod/autograft group, 2 in 4.5mm rod/allograft group, and 1 in 5.5mm rod/allograft group and as such the majority of rod fractures occurred early in the overall
Intervertebral fusion following anterior scoliosis surgery

TASF patient series. The surgical technique in our centre has been an evolution of steps in an attempt to provide the best possible surgical outcomes for patients and reduce the incidence of complications. When the rod fractures were identified early in our series of patients, the titanium rod diameter was increased from 4.5 to 5.5 mm and the bone graft changed from autograft (either rib head or iliac crest) to mulched femoral head allograft densely packed into well prepared disc spaces. So the rate of fracture in the earlier cases likely reflects that the quality of both the disc excision and bone graft packing of the disc spaces having improved with experience, which is also supported by other TASF literature [8, 10, 11]. Since changing both the graft material to allograft and increasing the rod size, the incidence of rod breakage has decreased dramatically in the larger series. There have been only three rod fractures found to date from the most recent 155 TASF cases since these change were made [6, 11, 15], such that the current complication rates for the full TASF series (210 patients with follow-up range 2.0 to 14.7 years) is 10.0% for rod fracture and 5.5% for screw related complications.

Radiographic pseudarthrosis has been shown to not necessarily correlate with final SRS questionnaire clinical outcome scores [21]. A recent study in 2010 [7] analysed patient satisfaction for a cohort of 100 patients from the TASF case series, with minimum 24 months follow up and found no significant differences in SRS-24 satisfaction scores between patients with rod fractures or screw-related complications compared to those without complications. The current study cohort SRS-24 results supports these previous findings, with overall mean SRS scores similar for patients with and without mechanical complications.

Anterior scoliosis correction using flexible rods has been associated with a higher complication rate when compared with posterior segmental instrumented fusions [37, 38]. Betz et al [39] found a 31% rod breakage rate compared with 1% for posterior instrumentation, but used smaller threaded rods
than those in the current study. Newton et al [4] reported a 7% rod fracture rate for a group of 41 TASF patients. Sweet et al [21] found rod diameter to be an important variable for pseudarthrosis and subsequent implant failure in anterior spinal fusion. In their group, all five patients with a pseudarthrosis had 5 or 5.5mm rods, whereas none of the patients with 4.5 mm rods developed this complication. The paper also suggested that because the larger rods are stiffer, they may prevent settling of endplates into a stable configuration and potentially favour a load-bearing situation leading to implant failure by exceeding the endurance limit of a single rod. Sweet et al explained that implant failure may allow further settling and compression at the non-union site, providing a favorable environment for fusion to occur finally. Supporting this theory, surgeons at our centre have found it rarely necessary for patients with a broken anterior rod to require revision surgery [6, 7], a finding which has been reported by other surgeons performing TASF [2, 4]. There has been no statistically significant increase in rib hump or in major, instrumented or compensatory curve Cobb angles after a rod fracture in the larger series [6].

Top screw pullout rates as high as 18% have been reported following anterior scoliosis correction procedures [40]. Partial proximal screw pullout occurred in five of the study patients and fusion scores at the IV space immediately below were lower compared to the levels without (1.35 versus 1.95) but not statistically significant. These cases were spread throughout the series and in most cases were noted either prior to discharge from hospital or at the first follow-up visit at two months after surgery. These patients were observed and did not require revision surgery for additional curve progression. Screw pullout can be related to patient bone density at the ends of the construct rather than the state of interbody fusion which does not come into play in the first weeks after surgery [40-42] when this occurs. So the significance of the trend for lower fusion scores at minimum two years after TASF surgery in this patient group is unclear.
A shortcoming of the current study is that there is a combination of factors at play that cannot be truly separated in the case of rod diameter versus the cessation of postoperative bracing. Subsequent experience at our centre has shown that patients do not need to be braced after TASF surgery to avoid rod breakage in the 5.5mm rod group of patients. However, this group was noted to have a lower fusion score than the 4.5mm rod group which may even suggest they require a smaller fusion mass to achieve a stable intervertebral segment with the larger diameter rod.

**Conclusion**

Rod diameter (larger), intervertebral level (proximal or distal relative to the apical levels), lateral position in disc space (further from rod) and the occurrence of a rod fracture all significantly reduce fusion scores, while graft type (autograft or allograft) does not affect scores. However, the assumed link between higher fusion score and better clinical outcome must be treated with caution, because in this series, 8 of the 11 rod fractures occurred in the 4.5mm titanium rod with autograft subgroup, even though 5.5mm rods have lower fusion scores and there was no overall difference in fusion scores between autograft and allograft. Thus it is suggested that with the use of the stiffer 5.5mm rod, less bony fusion mass may be required for a stable construct after anterior instrumented correction of progressive thoracic scoliosis. Taken together with previously published clinical results on this patient cohort [6], we propose that a full 100% radiologic fusion of the disc space is not necessary for successful clinical outcomes such that partial intervertebral fusion secures successful clinical outcomes in thoracoscopic scoliosis surgery.
Conflict of Interest

No financial support was received in support of this work. No benefits in any form have been or will be received from a commercial party related to the subject of this manuscript.

Acknowledgements

The authors thank Dr James Cordell-Smith for his assistance with fusion grading for this study.

Tables

**Table 1.** Analysis of intervertebral (IV) fusion scores for all subgroups with left para-sagittal being furthest from the rod, mid sagittal being midline of vertebrae and right para-sagittal being closest to the rod on the reformatted sagittal low dose CT images.
Table 2. Intra-observer and Inter-observer variability in fusion scores

<table>
<thead>
<tr>
<th>Intra-observer variability</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean signed intra-observer difference</td>
<td>-0.07</td>
</tr>
<tr>
<td>Mean unsigned intra-observer difference</td>
<td>0.35</td>
</tr>
<tr>
<td>Standard Deviation of intra-observer difference</td>
<td>0.55</td>
</tr>
<tr>
<td>95% Confidence Interval (1.96 × Standard Deviation)</td>
<td>1.08</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Inter-observer variability</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Deviation of difference between 2 observers (absolute)</td>
<td>0.58</td>
</tr>
<tr>
<td>Inter-observer Error (√2 × Standard Deviation)</td>
<td>0.82</td>
</tr>
<tr>
<td>95% Confidence Interval (1.96 × Standard Deviation)</td>
<td>1.61</td>
</tr>
</tbody>
</table>

Table 3. SRS-24 mean total score and mean score for individual domains (± SD) at minimum 24 months after surgical correction for the following groups; all patients (n=42) and subgroups no mechanical complications (n=25), rod fracture (n=11), top screw pullout (n=6) groups.

<table>
<thead>
<tr>
<th>SRS-24</th>
<th>All patients</th>
<th>No mechanical complications</th>
<th>Rod fracture</th>
<th>Screw pullout</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total (24 questions / 120)</td>
<td>98.9 ± 7.3</td>
<td>98.0 ± 7.3</td>
<td>101.5 ± 8.2</td>
<td>98.0 ± 6.2</td>
</tr>
<tr>
<td>Mean total score / 5</td>
<td>4.1 ± 0.3</td>
<td>4.1 ± 0.3</td>
<td>4.2 ± 0.3</td>
<td>4.1 ± 0.3</td>
</tr>
<tr>
<td>Domain Scores / 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td>4.3 ± 0.5</td>
<td>4.3 ± 0.5</td>
<td>4.3 ± 0.6</td>
<td>4.2 ± 0.5</td>
</tr>
<tr>
<td>General self-image</td>
<td>4.0 ± 0.7</td>
<td>4.1 ± 0.7</td>
<td>4.0 ± 0.7</td>
<td>3.8 ± 0.6</td>
</tr>
<tr>
<td>General function</td>
<td>4.2 ± 0.3</td>
<td>4.3 ± 0.2</td>
<td>4.3 ± 0.2</td>
<td>4.3 ± 0.2</td>
</tr>
<tr>
<td>Activity level</td>
<td>4.5 ± 0.7</td>
<td>4.5 ± 0.7</td>
<td>4.6 ± 0.4</td>
<td>4.5 ± 0.8</td>
</tr>
<tr>
<td>Postoperative self-image</td>
<td>3.5 ± 0.5</td>
<td>3.4 ± 0.5</td>
<td>3.8 ± 0.5</td>
<td>3.7 ± 0.6</td>
</tr>
<tr>
<td>Postoperative function</td>
<td>3.2 ± 0.8</td>
<td>3.0 ± 0.7</td>
<td>3.7 ± 0.8</td>
<td>3.7 ± 0.8</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>4.5 ± 0.5</td>
<td>4.6 ± 0.4</td>
<td>4.5 ± 0.6</td>
<td>4.2 ± 0.9</td>
</tr>
</tbody>
</table>
Figures

Figure 1. Preoperative and 2 year postoperative standing radiographs (A, B) and photographs (C, D) of patient following thoracoscopic anterior spinal fusion surgery.

Figure 2. Illustrates the location of the three reformatted sagittal images used in the study; A) Right para-sag (closest to the rod), B) Mid Sag (middle of the IV space), and C) Left para-sag (furthest from the rod).
**Figure 3.** Sucato fusion grading method [29] for scoring the area of the IV space where bone is visible. Grade 0 no fusion, Grade 1 < 50% area, Grade 2 between 50-75%, Grade 3 >75%, Grade 4 completely across the area.

**Figure 4.** Example of a A) top screw pullout and B) rod fracture shown on coronal low dose CT images.
Intervertebral fusion following anterior scoliosis surgery

Figure 5. Graph showing the mean fusion scores for the various rod diameters/stiffness types for all three sagittal reconstructed planes and an overall fusion score when all three sagittal planes were combined. Note that these rod comparisons were for cases that used allograft, and Right para-sag is closest to the rod.
Figure 6. Graph showing the mean fusion scores for the different graft types when 4.5mm Titanium Rod was used, showing all three sagittal reconstructed planes and an overall fusion score when all three planes were combined. Note: R para-sag is closest to the rod.
**Figure 7.** Graph showing the mean fusion scores for the IV levels fused in TASF with Level 1 being the most cephalad level and Level 6 being the most caudal level, for all three sagittal reconstructed planes. Note: R para-sag is closest to the rod.
Figure 8. Graph showing the mean fusion score for the levels where a rod fractured versus the levels where there was no rod fracture for all three sagittal reconstructed planes and an overall fusion score when all three planes were combined. Note: R para-sag is closest to the rod.
Intervertebral fusion following anterior scoliosis surgery

References

Intervertebral fusion following anterior scoliosis surgery


