The international spine registry SPINE TANGO: status quo and first results

Markus Melloh · Lukas Staub · Emin Aghayev · Thomas Zweig · Thomas Barz · Jean-Claude Theis · Albert Chavanne · Dieter Grob · Max Aebi · Christoph Roeder

Received: 28 July 2006 / Revised: 12 March 2008 / Accepted: 30 March 2008 © Springer-Verlag 2008

Abstract With an official life time of over 5 years, Spine Tango can meanwhile be considered the first international spine registry. In this paper we present an overview of frequency statistics of Spine Tango for demonstrating the genesis of questionnaire development and the constantly increasing activity in the registry. Results from two exemplar studies serve for showing concepts of data analysis applied to a spine registry. Between 2002 and 2006, about 6,000 datasets were submitted by 25 centres. Descriptive analyses were performed for demographic, surgical and follow-up data of three generations of the Spine Tango surgery and follow-up forms. The two exemplar studies used multiple linear regression models to identify potential predictor variables for the occurrence of dura lesions in posterior spinal fusion, and to evaluate which covariates influenced the length of hospital stay. Over the study period there was a rise in median patient age from 52.3 to 58.6 years in the Spine Tango data pool and an increasing percentage of degenerative diseases as main pathology from 59.9 to 71.4%. Posterior decompression was the most frequent surgical measure. About one-third of all patients had documented follow-ups. The complication rate remained below 10%. The exemplar studies identified “centre of intervention” and “number of segments of fusion” as predictors of the occurrence of dura lesions in posterior spinal fusion surgery. Length of hospital stay among patients with posterior fusion was significantly influenced by “centre of intervention”, “surgeon credentials”, “number of segments of fusion”, “age group” and “sex”. Data analysis from Spine Tango is possible but complicated by the incompatibility of questionnaire generations 1 and 2 with the more recent generation 3. Although descriptive and also analytic studies at evidence level 2++ can be performed, findings cannot yet be generalised to any specific country or patient population. Current limitations of Spine Tango include the low number and short duration of follow-ups and the lack of sufficiently detailed patient data on subgroup levels. Although the number of participants is steadily growing, no country is yet represented with a sufficient number of hospitals. Nevertheless, the benefits of the project for the whole spine community become increasingly visible.

Keywords Spine registry · Posterior spinal fusion · Dura lesion · Length of hospital stay · Predictors
Introduction

Spine Tango is the first international spine registry, with 6,000 submitted datasets from 25 centres in and outside Europe [1, 19]. Although content, set-up and workflow have been presented previously, results of Spine Tango have not yet been published [17]. The Swedish spine registry has already shown that a national registry can generate valid and meaningful data [6].

With Spine Tango generation 3 now implemented, this article compares the main developmental stages of Spine Tango regarding its content, with focus on demographics, surgery and follow-up data. Possibilities of generating scientific evidence from an international spine registry are discussed. As a proof of concept of data analysis, two relevant examples about predictors of dura lesions and length of hospital stay in posterior spinal fusion are presented [16].

Genesis of Spine Tango

In the mid 1990s, the head of the spine unit at the Schulthess Clinic Zurich, Prof Dieter Grob, developed a documentation system for quality assessment in spinal surgery. At EuroSpine 2000 in Antwerp, a web-based prototype of this documentation system named Spine Tango was introduced. It had been researched and developed by the working group of the revived Maurice E. Müller Center for Education and Documentation (MEM-CED) under the supervision of Prof Max Aebi. One year later the first beta-test version was presented at the EuroSpine in Gothenburg, and in January 2002 the second beta-test version was shown on the Orthoglobe portal.

In November 2002, Spine Tango generation 1 started with six pioneer hospitals submitting data online to the central database in Berne, Switzerland. This Spine Tango generation ran for eight months before it was replaced by a generation 2 in July 2003 because some deficits in structure and content had become obvious.

Twenty-two months later, in May 2005, generation 3 was released. At this stage, both, an online and paper-based Spine Tango documentation system were available for the first time. Forms were drastically reduced in length, which increased acceptance and the number of participants. Parallel to Spine Tango generation 3, patient self-assessment forms were introduced and officially recommended by the Spine Society of Europe (SSE) [3, 12]. In 2006 a second and streamlined version of Spine Tango generation 3 was worked up. This time, however, compatibility with the previous version was assured, so that another disruption in data continuity was avoided. Moreover, the Spine Tango committee considers the currently used Spine Tango forms as a final version, which can be used over the next 5–10 years.

Current situation of Spine Tango

At present 25 centres in nine countries submit data to the Spine Tango registry. Considering all candidates that are in their rollout phase, the registry could soon expand to include data from about 52 hospitals in 18 countries.

To cope with administrative issues and the legal requirements of data anonymisation, national Spine Tango modules are inevitable [17, 18]. A module is an interlinked web server that acts as a filter between the user and the central database for the purpose of encoding, that is anonymising all sensitive patient data and also that of the users. The central database solely hosts clinical data without patient names, addresses or the original medical record number of the treating institution (Fig. 1) [18]. A main strategic goal of Spine Tango is a further expansion of the network of clinics to achieve an even broader coverage of countries and institutions. In the future, at least one Spine Tango module should be installed in every participating country. In countries with many participating clinics (e.g. Germany) or with associated medical fields (orthopaedic surgery, neurosurgery) the installation of several modules might become necessary.

Indisputably, a marketing concept is needed for the Spine Tango registry. An acquisition of new centres via national spine societies seems an obvious strategic approach. While this article focuses on summarising initial demographic, surgical and follow-up data, future analyses will also include patient-based outcomes. The patient self-assessment questionnaires are currently being implemented and validated in various languages.

Four national Spine Tango modules are in operation to date: the Swiss/International module at the MEM Research Center for Orthopaedic Surgery, University of Berne (formerly MEM-CED), for all Swiss users and all international participants without an own national module; the Austrian module at the Orthopaedic Hospital Speising, Vienna; the German module at the Department of Orthopaedic Surgery, University Hospital, Cologne; and the US module located in Ormond Beach, FL. Another three modules are in the process of rollout in Italy, Brazil and southern Germany where a new cluster of hospitals led to the need of a second German Spine Tango module.

The user community (including candidates that are in their rollout phase) is made up of 41 Departments of Orthopaedic Surgery and 11 Departments of Neurosurgery. One-fourth of these clinics (13/52) are University Hospitals, which are destined to guide the Spine Tango registry as opinion leaders.
Germany has the largest and fastest growing user community. The German example shows that an organisational concept is needed to run Spine Tango successfully. Looking at the location of the German Spine Tango clinics, several clusters have formed. They emphasise the necessity of a division of Germany into regional administrative subcentres. The future will show whether other countries will have comparable cluster formation.

**Methods/patients**

From November 2002 to November 2006 almost 6,000, patient datasets were submitted (generation 1, 831; generation 2, 1,554; generation 3, 3,437). Looking at the dates of surgery, patients of the first generation were operated between January 2002 and July 2003, patients of the second generation between January 2002 and April 2005, and those of the third one between March 2003 and November 2006. This shows a certain time lag of data entry.

Twenty-six percent of the patients in generation 1, 6% in generation 2, and 20% in generation 3 had their intervention in a time period before the creation date of the respective dataset in the Spine Tango database. This means that the data were entered retrospectively. In contrast, about 80% of the data were entered in real time. The increasing ratio can be explained with the new possibility of paper-based documentation from generation 3 onwards.

Because data collection is organised in a prospective mode in most centres and only the final submission of data occurred with a delay, data validity was not affected. The creation date was limited by the runtime of each Spine Tango version so that there was no transitional period. With the start of each new form generation the previous one was shut down, but pending cases could also be completed thereafter.

The average number of patients submitted monthly was 104 per month for generation 1, 71 per month for generation 2 and 181 per month for generation 3.

Due to changes in number and scale of most items, frequency statistics had to be calculated separately for Spine Tango generations 1, 2 and 3. All values are described as frequencies and percentages, with the exception of age and length of hospital stay for which medians and interquartile ranges were used.

Two exemplar studies on data of Spine Tango generation 3 were performed. The first assessed factors predicting the occurrence of dura lesions in posterior spinal fusion using a generalised linear model. The second study looked for predictors of length of hospital stay. Age, sex, main pathology, number of spinal segments of posterior fusion, level of fusion, number of previous spinal surgeries, operation time, centre of intervention (including annual number of fusions undertaken and academic status), surgeon credentials and type of fusion (fusion vs. fusion plus instrumentation vs. fusion plus instrumentation and cage implantation) were included in the models. The level of significance was set to 0.05 throughout the study. All statistical analyses were conducted using SAS 9.1 (SAS Institute Inc, Cary, NC).

**Results**

Demographics and surgery data

Table 1 compares the patient demographics and frequencies of selected Spine Tango variables in generations 1, 2 and 3. These data focus on main pathology, additional
Table 1  Demographics and variables of primary datasets in Spine Tango generations 1, 2, and 3

<table>
<thead>
<tr>
<th></th>
<th>Spine Tango generation 1</th>
<th>Spine Tango generation 2</th>
<th>Spine Tango generation 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of primary datasets</td>
<td>n = 831</td>
<td>n = 1554</td>
<td>n = 3437</td>
</tr>
<tr>
<td>Demographics</td>
<td>52.3 years (IQR: 37.5–65.6)</td>
<td>56.0 years (IQR: 43.3–9.1)</td>
<td>58.6 years (IQR: 44.3–70.6)</td>
</tr>
<tr>
<td>Sex (%)</td>
<td>Females 53.3</td>
<td>Females 55.1</td>
<td>Females 53.8</td>
</tr>
<tr>
<td>Main pathology</td>
<td>Degenerative disease 59.9</td>
<td>Degenerative disease 72.7</td>
<td>Degenerative disease 71.4</td>
</tr>
<tr>
<td>Second most freq. main pathology (%)</td>
<td>Deformity 9.0</td>
<td>Deformity 5.7</td>
<td>Spondylolisthesis 8.4</td>
</tr>
<tr>
<td>Third most freq. main pathology (%)</td>
<td>Trauma/fracture 7.8</td>
<td>Spondylolisthesis 5.5</td>
<td>Deformity 6.6</td>
</tr>
<tr>
<td>Additional pathology</td>
<td>None (%) 66.8</td>
<td>78.2</td>
<td>68.1</td>
</tr>
<tr>
<td>Surgical goals</td>
<td>Neurodecompression 59.8</td>
<td>Neurodecompression 66.4</td>
<td>Pain relief 92.6</td>
</tr>
<tr>
<td>Level of procedure</td>
<td>Lumbar 35.7</td>
<td>Lumbar 47.9</td>
<td>Lumbar 55.0</td>
</tr>
<tr>
<td>Surgical measures</td>
<td>Posterior decompression 40.0</td>
<td>Posterior decompression 36.9</td>
<td>Posterior decompression 71.9</td>
</tr>
<tr>
<td>Length of hospital stay</td>
<td>7 days (IQR: 4–10)</td>
<td>7 days (IQR: 4–10)</td>
<td>8 days (IQR: 5–13)</td>
</tr>
</tbody>
</table>

pathology, surgical goals, level of procedure, surgical measures and length of hospital stay. Because the content of Spine Tango varies between the three generations due to the process of evolution in practice, only comparable variables were selected.

Between 2002 and 2006 there was an age increase from 52.3 to 58.6 years in the data pool. Degenerative disease was the predominant main pathology in all versions of Spine Tango, with a rise from 59.9 to 71.4%. The lumbar spine as most frequent level of procedure increased from 35.7 to 55.0%. The lumbar-sacral spine was approached less often, declining from 27.3 to 22.0%. The commanding surgical measure in all versions of Spine Tango was the sole posterior decompression, but over the years posterior stabilisation and posterior fusion gained ground.

Females dominated among spine patients in the registry. Additional pathology was not specified in about one-third of all cases. Answers given for additional pathology were heterogeneous, with degenerative disease, osteoporosis,
deformity and spondylolisthesis being the most frequent. Cervical procedures were documented less frequently, with mid-lower cervical procedures being the largest subgroup. Posterior surgical measures were recorded much more often than anterior ones by the Spine Tango community. Average length of hospital stay varied between 7 and 8 days in the three generations.

A closer look at the main pathology revealed the dominating position of degenerative diseases in the Spine Tango data pool (Fig. 2). All other main pathologies had a proportion of less than 10% each.

Fig. 2 Frequency of main pathology in Spine Tango generations 1, 2 and 3 in percent

Follow-up data

Totally 2,083 patients had follow-ups. The follow-up rate of the three generations varied between 23 and 38.3% (Table 2). Median follow-up time in generation 1 was 8.7 weeks, 12.1 weeks in generation 2 and 8.6 weeks in generation 3. The comparable variables of generations 1 to 3 are the achievement of surgical goals, work status and rehabilitation. Overall outcome from the surgeon’s point of view was only introduced in generation 3 and because of its importance in the follow-up process it deserves mention here. One-third of all surgeons estimated their patients’ overall outcome as excellent. Only 5% judged it as poor.

Complications at follow-up

At the time of follow-up, less than 10% of all patients had reported complications. Sensory disturbance was the most frequent complication in all three Spine Tango generations. Nevertheless, it steadily decreased over time in the data pool.

The various types of complications with frequencies below 5% were widely scattered in all Spine Tango generations (Fig. 3).

Table 2 Follow-up data in Spine Tango generation 1, 2, and 3

<table>
<thead>
<tr>
<th>Achievement of surgical goals</th>
<th>Spine Tango generation 1</th>
<th>Spine Tango generation 2</th>
<th>Spine Tango generation 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Achieved (%)</td>
<td>80.6</td>
<td>93.3^1</td>
<td>66.2^1</td>
</tr>
<tr>
<td>Partially achieved (%)</td>
<td>14.7</td>
<td>25.5^1</td>
<td>41.3^1</td>
</tr>
<tr>
<td>Not achieved (%)</td>
<td>4.7</td>
<td>10.3^1</td>
<td>8.1^1</td>
</tr>
<tr>
<td>Work status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not at work since OP (%)</td>
<td>25.1</td>
<td>22.9</td>
<td>19.0</td>
</tr>
<tr>
<td>Started partially same job (%)</td>
<td>14.7</td>
<td>8.2</td>
<td>9.9</td>
</tr>
<tr>
<td>Fully reintegrated (%)</td>
<td>15.7</td>
<td>16.6</td>
<td>18.7</td>
</tr>
<tr>
<td>Resumed work, quit again (%)</td>
<td>0</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>Resumed work, different job (%)</td>
<td>0</td>
<td>0</td>
<td>0.4</td>
</tr>
<tr>
<td>Dismissed (%)</td>
<td>0.5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Retired (%)</td>
<td>38.7</td>
<td>49.1</td>
<td>41.4</td>
</tr>
<tr>
<td>Housewife, student, or child (%)</td>
<td>5.2</td>
<td>2.0</td>
<td>16.1</td>
</tr>
<tr>
<td>Other (%)</td>
<td>0</td>
<td>0</td>
<td>1.1</td>
</tr>
<tr>
<td>Overall outcome (surgeon)</td>
<td></td>
<td></td>
<td>36.2</td>
</tr>
<tr>
<td>Excellent</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Good</td>
<td>–</td>
<td>–</td>
<td>42.9</td>
</tr>
<tr>
<td>Fair</td>
<td>–</td>
<td>–</td>
<td>15.1</td>
</tr>
<tr>
<td>Poor</td>
<td>–</td>
<td>–</td>
<td>5.1</td>
</tr>
<tr>
<td>Not applicable</td>
<td>–</td>
<td>–</td>
<td>0.8</td>
</tr>
</tbody>
</table>

^1 Multiple choice question: answers add up to >100%
Exemplar studies on Spine Tango data

In order to show how a methodologically sound analysis can be conducted on an observational registry data set, we conducted two independent analytical studies on data from generation 3 only. It was chosen because it had the highest data consistency and number of observations.

The first study examined predictors for the occurrence of dura lesions during posterior spinal fusion. Dura lesions were the most frequent type of intra-operative complication in the registry, and thus were chosen as main outcome variable [21].

Between May 2005 and November 2006, data of 3,437 patients were documented in Spine Tango generation 3. Nine hundred and twenty-nine patients who had been treated with posterior spinal fusion were included in this study. Clinics contributing fewer than 25 cases with posterior spinal fusion were excluded from the analysis. Median age was 62.7 years (min 12.5 years, max 90.5 years) with a female-to-male ratio of 2:1. In 18 of 929 cases a dura lesion was documented (1.9%). The prevalence differed significantly among clinics ranging from 0.6 to 8.5%.

A multiple logistic regression model was built with the following potential predictor variables for the occurrence of dura lesions: age, sex, main pathology, number of spinal segments of posterior fusion, level of fusion, number of previous spinal surgeries, operation time, centre of intervention (including number of fusions in the study period and academic status), surgeon credentials and type of fusion (fusion vs. fusion plus instrumentation vs. fusion plus instrumentation and cage implantation).

Centre of intervention \((p = 0.02)\) and number of fused segments \((p = 0.018)\) were revealed as predictors of the occurrence of dura lesions in posterior spinal fusion. Number of fusions per centre (min 25, max 526) and its academic status had no influence on the rate of dura lesions. Fusions of four or more segments showed a dura lesion rate that was three times higher than fusions of less than four segments (Fig. 4). No other covariate showed an influence on the rate of dura lesions.

In the second study, multivariate linear regression was applied in order to find predictors that influence length of hospital stay in patients with posterior spinal fusion. The length of hospitalisation varies widely in patients with posterior spinal fusion, and even more if a posterior or anterior approach was used [5, 7, 13, 14].

Seven hundred and ninety patients with degenerative disease (614 patients) or spondylolisthesis (176 patients) who had been treated with posterior decompression and fusion were included in this study. Median age was 62.8 years (min 13.2 years, max 89.8 years), with a female-to-male ratio of 2:1. Median length of hospital stay was 11 days (IQR 8–14 days). Length of hospital stay of more than 100 days was considered as erroneous data entry and cases were excluded.

Centre of intervention \((p < 0.0001)\) was revealed as a highly significant predictor of length of hospital stay (min 7 days, max 14 days; Fig. 5). Number of fusions per clinic (min 25, max 434) and academic status of clinic had no influence on length of hospital stay. Further significant predictors were surgeon credentials (surgeons in training, 8.5 days; specialised spine surgeons, 11 days; orthopaedic or neurosurgeons, 12 days; \(p = 0.001\)), number of spinal segments of posterior fusion (1 segment, 10 days; 2–3 segments, 12 days; 4–5 segments, 12.5 days; >5 segments, 15 days; \(p = 0.002\)) and age group (<50 years, 9 days; 50–59 years, 11 days; 60–69 years, 12 days; ≥70 years, 13 days; \(p = 0.01\)). Borderline significance was found for sex (women, 12 days; men, 10 days; \(p = 0.05\)). All other variables showed no influence on length of hospital stay.
Conclusions

Spine Tango has achieved a firm position as international spine registry and with its increasing acceptance it is also gaining importance. Currently there are more than four dozen participating clinics in an active status or in an implementation phase. The strengths of Spine Tango include a potentially very large network, the participation of a specialised international society and an academic partner with expertise and extensive experience in registry implementation.

Spine Tango’s future potential lies in cooperation with other spine registries and the administration by national spine societies. The large knowledge gained in registries like the Swedish lumbar spine register could certainly be advantageous for solving problems that the Tango will only face in the future.

The descriptive analysis of Spine Tango demographics, surgery and follow-up data are not yet representative. Similarly, the increasing number of participating clinics over the course of Spine Tango’s genesis makes prohibitive a generalisation of observed trends in demographic data. At this point in time, they must be understood as inventory description of the database that nicely depicts the evolution and growth of the data pool. In the near future the respective national participants may form a group large enough to be a representative sample of hospitals with a similarly representative patient population.

The analysis of observational registry data with appropriate statistical methods can generate findings with an evidence level 2++. It is slowly recognised that RCTs have limitations [20] and may be unnecessary, inappropriate, impossible to conduct, or inadequate [17]. Therefore, observational studies, with a lower level of evidence but higher feasibility and external validity are increasingly applied [2, 17].

The 1.9% prevalence of dura lesions in our study was in line with literature reports of prevalence that range between 0.3 and 29% [9, 11, 15, 21].

Our main conclusion was that “centre of intervention” and “number of segments of fusion” are predictors of dura lesions in posterior spinal fusion. Unanticipated, the number of spinal surgeries and academic status of the clinic had no influence on the occurrence of dura lesions. Therefore, a subgroup analysis should be performed assessing further covariates. However, this goes beyond our current possibilities since the absolute number of dura lesions does not allow further stratification, and only a limited number of variables can be recorded in a spine registry. Furthermore, the large variation of the occurrence of dura lesions among the included clinics may indicate different opinions about when a dura lesion must be considered a complication or not. Hence, they may not be documented the same way in all participating centres and conclusions about the safety of posterior fusion in the respective hospitals are problematic.

Surgeon based reports of complications are a general weakness of medical registries and an obvious solution for that problem is demonstrated by the Swedish lumbar spine register—interviewing patients about re-interventions in the same or another hospital because of an adverse event from the initial surgery. As long as anonymous data is presented, “centre of intervention” as predictor of the occurrence of dura lesions does not have any implication for surgical practice. However, the information can serve as orientation displayed in quarterly reports that are sent to each participating centre. If the surgeons conclude that their definition of dura lesion is too strict or their frequency of dura lesions too high remains open and depends on the individual case. Nevertheless, the benchmarking possibilities certainly initiate a thought process and eventually even a change of “behaviour”, be it in the theatre or on the Tango questionnaires. On the other hand, knowing that in posterior fusions of four or more segments a threefold higher risk of dura lesions exists may have an impact on the surgeon’s decision about the number of segments to be included or on his or her carefulness during surgery.

The 11 days median length of hospital stay in the second analysis seemed quite high. This can be explained by the more invasive character of the included interventions compared to the pool data, where sole decompression surgery, total disc replacement and percutaneous procedures can be found. A subgroup analysis would be needed to differentiate between conventional and minimally invasive fusion procedures, as well as between different main pathologies that may result in diverse length of hospital stay [4, 8, 13].

“Centre of intervention”, “surgeon credentials”, “number of segments of fusion”, “age group” and “sex” were predictors of length of hospitalisation in posterior spinal fusion.

Fig. 5 Length of hospital stay in posterior spinal fusion in seven selected clinics
spinal fusion. These results are in accordance with the literature showing a positive correlation between number of segments fused and age on one side, and length of hospital stay on the other side [10, 22]. A shorter length of stay in surgeries performed by surgeons in training can be explained by the lower number of segments fused in these procedures. A longer length of stay in surgeries performed by surgeons in training can be explained by the lower number of segments fused in these procedures. A longer length of stay in surgeries performed by surgeons in training can be explained by the lower number of segments fused in these procedures.

There are some obvious and indisputable limitations to the analysis of the Spine Tango data pool at this point in time. There is a low coverage of documented interventions within departments and an even lower one on a national or international level. Bias in reporting complications is immanent for various reasons. As long as participation in registries is conducted on a voluntary basis, low participation rates in general and a rather careful reporting of complications and revisions of those who participate are an almost natural behaviour. Only if everybody has to document according to binding rules and with monitoring mechanisms will the participation rate and representativeness of results increase. Governmentally monitored registries may result in a nationwide coverage of up to 99%, as is the case of the National Joint Register in New Zealand. Nevertheless, voluntary registries can also be powerful instruments for generation of hypotheses that can then be investigated in a prospective multicentre study setting using the registry infrastructure and its most motivated participants.

**Outlook**

Despite various partially incompatible generations the Spine Tango data pool nicely depicts the laborious genesis and the fact that now first analyses are possible and result in clinically meaningful findings. These can, however, not yet be generalised for day to day practise but rather serve as orientation for the participants and indication of the future potential of Spine Tango.

Since there is stability in content since the year 2005, increasing user quantity and activity and additional value in terms of patient based outcome assessment, future data evaluations should reveal even more relevant findings with a higher external validity. This will make the impact of Spine Tango more obvious and have a positive effect on acquisition of new participants.

The newly implemented features on the system like improved online statistics with direct benchmarking mechanisms will additionally support the success of the European Spine registry, which is already taking on global dimensions.

**References**

19. Springer