Inter-rater reliability of the tracking and reporting outcomes of procedural sedation tool

Aaron Conway PhD1,2,3,✉, and Katina Corones-Watkins PhD4

1 Peter Munk Cardiac Centre, University Health Network, Toronto, Canada  
2 Lawrence S. Bloomberg Faculty of Nursing, University of Toronto, Toronto, Canada  
3 School of Nursing, Queensland University of Technology (QUT), Brisbane, Australia  
4 School of Nursing, Giffith University, Brisbane, Australia

✉ Correspondence: [Aaron Conway PhD <[aaron.conway@utoronto.ca](mailto:aaron.conway@utoronto.ca)> | +1 416 946 7112 | 585 University Ave | Toronto, ON M5G 2N2](mailto:aaron.conway@utoronto.ca)

## Abstract

Clinical guidelines recommend that outcomes of procedural sedation are audited. The Tracking and Reporting Outcomes of Procedural Sedation (TROOPS) tool, developed by the International Consortium for the Advancement of Procedural Sedation, is intended to be suitable for use in all locations by sedation providers. We evaluated the feasibility and reliability of using the TROOPS tool for sedation that is administered by nurses in the cardiac catheterization laboratory. The version of TROOPS we used contained items in the minor, intermediate and sentinel categories. Two nurses independently completed the TROOPS tool for 40 patients who underwent a procedure with sedation. A sedation-related adverse event was selected by at least one of the raters for 21 of the patients (52%; 95% CI = 0.38 to 0.67). Most were minor events related to the airway and breathing category of the TROOPS tool (n = 17; 42%; 95% CI = 0.29 to 0.58). The remaining were intermediate-severity events related to sedation quality. No events for the other categories of the TROOPS tool were selected. The intra-class correlation coefficient between paired nurse ratings of the TROOPS tool was 0.78 (95% CI = 0.43 to 0.92). Most nurses (85%, 95% CI = 73% to 92%) reported that it took less than 1 minute to complete the tool. In summary, paired ratings of sedation-related adverse events by nurses achieved moderate consistency using a version of the TROOPS tool containing items from the minor category of the airway and breathing domain. Use of this tool in clinical practice seems feasible.

**Keywords**

Procedural sedation and analgesia; conscious sedation; nursing; patient safety; anesthesia

## Introduction

The administration of sedative and analgesic medications during medical procedures is associated with several unintended effects with the potential to compromise patient safety. However, if systems are in place to manage these complications, serious sedation-related serious adverse events are preventable. One central aspect of ensuring patient safety during sedation that is recommended in clinical guidelines is auditing of sedation-related outcomes.[1–5] For example, the Practice Guidelines for Moderate Procedural Sedation and Analgesia from the American Society of Anesthesiologists Task Force in Moderate Procedural Sedation recommend the creation and implementation of a quality improvement process to measure adverse events and unsatisfactory sedation.[2] Likewise, the Canadian Anesthesiologists Society recommends that a specific mechanism must be established for auditing procedural sedation.[1]

The Tracking and Reporting Outcomes of Procedural Sedation (TROOPS) tool was developed by the International Consortium for the Advancement of Procedural Sedation.[6] The tool is intended to be suitable for use in all locations by sedation providers regardless of the particular discipline. There are two versions of the tool. One is intended to be used as a strictly quality improvement tool and the other in research. The quality improvement version includes events categorized as intermediate and sentinel. Intermediate items in the TROOPS tool are defined as having the *potential* to endanger patients if not properly managed or reflecting suboptimal sedation quality. These items require timely reporting. Sentinel adverse events are defined as being life-threatening and warrant immediate reporting and the highest level of peer scrutiny.

The research tool includes additional *minor* adverse events, which are defined as those that pose little risk given appropriate provider skills and monitoring.[6] Using the intermediate and sentinel event items included in the TROOPS tool for quality improvement seems appropriate due to the rigorous method in which it was developed. Although intended for a different purpose, specific items in the research version of the tool designated as *minor* may also be useful to track for the purposes of quality improvement because these events are more common. However, the feasibility and reliability of using the TROOPS tool for measuring the outcomes of sedation that is administered by nurses has yet to be evaluated. Such evaluation is required in order for the auditing component to be considered valid enough to inform practice improvement initiatives. For this reason, the aims of this study were to determine the:

1. Inter-rater reliability of nurses’ TROOPS ratings; and
2. Time taken for nurses to complete the TROOPS tool (calculated to determine the impact that integrating auditing of sedation outcomes into routine clinical practice would have on nurses’ workload).

## Methods

A prospective study in an urban private hospital in Australia was undertaken. The study was approved by the human research ethics committee (UCH HREC 2016.23.201).

### Participants

Patients undergoing procedures with nurse-administered sedation in a cardiac catheterization laboratory were eligible. A convenience sample was used.

### Measurements

A data collection form that contained three parts was developed to be used in this study. The first part contained items to collect demographic (age, sex) and clinical information (type of procedure, procedure start and end time, baseline oxygen supplementation and total doses of sedation used) about the included patients.

The second part consisted of the quality improvement version of the TROOPS tool, with the addition of minor adverse events from the ‘airway and breathing’ as well as ‘gastrointestinal’ domains from the research version of the tool.[6] Minor events from the ‘airway and breathing’ domain used were ‘increased or added supplemental oxygen,’ ‘airway repositioning’ and ‘tactile stimulation.’ Based on consultation with members of the clinical team, we elected not to include the minor items related to hypersalivation (suctioning for hypersalivation and anticholinergic for hypersalivation) because the medications used for sedation in the department (fentanyl and midazolam) are not associated with this particular adverse effect. Both minor events from the gastrointestinal domain of the research version were used. These were administration of anti-emetic and suction for emesis. The minor items from the allergy and neurological domains were not included because these adverse effects were not considered to occur frequently enough in this particular sedation context to warrant inclusion as a measure for quality improvement. Other modifications were made to the TROOPS tool through consultation with members of the clinical team in order to make it suitable for use by nurses in that particular department. Modifications included the removal of outcomes that were not feasibly assessed by nurses in the immediate post-procedure period, including permanent neurological deficit from the sentinel category of the neurological domain and pulmonary aspiration syndrome from the sentinel adverse events category of the airway and breathing domain.

The third part of the data collection form included one item for nurses to rate how long it took to complete the TROOPS tool ratings. Possible selections were ‘less than 1 minute,’ ‘between 1 and 3 minutes,’ ‘between 3 and 5 minutes,’ and ‘more than 5 minutes.’

### Data collection

Two nurses who performed the ‘scout’ and ‘monitor’ role during procedures performed with nurse-administered sedation completed the TROOPS tool. Ratings were performed independently as soon as possible after the completion of the procedure. Prior to commencement of the project, all nurses involved in administering sedation in the clinical department were invited to attend in-services where information about the project and the assessment tool was provided. We did not collect information on the number of nurses attending the education sessions.

### Data analysis

Inter-rater reliability was assessed by calculating the random effects one-way intra-class correlation between nurses’ paired ratings. Lower and upper confidence(95%) bounds associated with the intra-class correlation under the one-factor ANOVA model was used because each patient could have been rated by a different group of raters (i.e. nurses).[7] Feasibility was examined by calculating proportions for nurses’ selections about the time taken to complete the tool. Analyses were undertaken using R.[8] Data, statistical code and a computational environment required to reproduce the results is available.[insert ref]

## Results

Data were collected from January to November, 2017. Two nurses independently completed the TROOPS tool for 40 patients who underwent a procedure with sedation. A summary of the patient characteristics is presented in Table . Most patients were undergoing a cardiac implantable electronic device implant (e.g. pacemaker or implantable cardioverter defibrillator) or electrophysiology procedure. A combination of midazolam and fentanyl was used for all patients.

A sedation-related adverse event was selected by at least one of the raters for 21 of the patients (52%; 95% CI = 0.38 to 0.67). Most were minor events related to the airway and breathing category of the TROOPS tool (n = 17; 42%; 95% CI = 0.29 to 0.58). The remaining were intermediate-severity events related to sedation quality and patient experience. No events for the other categories of the TROOPS tool were selected. The intra-class correlation coefficient between paired nurse ratings of the TROOPS tool was 0.78 (95% CI = 0.43 to 0.92). Responses from fifty-one (85%, 95% CI = 73% to 92%) nurses indicated that it took less than 1 minute to complete the TROOPS tool. The remaining nurses responded that it took between 1 and 3 minutes to complete. There were 20 missing responses to this item for the 40 patients.

## Discussion

Reliability is a vital property of a measurement tool if results are intended to inform quality improvement. There was moderate consistency between paired TROOPS ratings. However, it should be noted that confidence intervals for the intra-class correlation were wide, spanning from poor (0.43) to very high levels of consistency (0.92). Additional research with a larger sample size would increase confidence in estimates of inter-rater reliability for the TROOPS tool.

Further research would be valuable to identify strategies that increase the reliability of using the TROOPS tool for rating sedation-related adverse events in practice. Pending such research, there are several recommendations for practice that can reasonably be drawn from these results. First, our study involved a fairly passive implementation process, which was limited to the provision of information. Instructions about how to use the tool were provided to nurses during scheduled in-service sessions. A more ‘active’ implementation process, including, for example, practical demonstrations of ‘correct’ ratings, may be needed to improve the reliability of TROOPS tool ratings. Video recordings of sedation encounters have been used to validate sedation scales in prior research.[9] Using videos of sedation encounters could be a practical and efficient approach to facilitate demonstrations of the correct usage of the TROOPS tool for rating sedation-related adverse events. Alongside education and instruction on the correct usage of the TROOPS tool, periodic evaluation of inter-rater reliability should be undertaken by procedural sedation programs electing to use the TROOPS tool with items from the minor category of the airway and breathing domain.

In our study, both nurses who completed TROOPS ratings for each patient were involved in the procedure, but they performed different roles. In the department in which data collection was performed, nurses performed the roles of ‘scout’ nurse and ‘monitor’ nurse. This is commonplace in the cardiac catheterization laboratory setting.[10] Based on the authors’ experiences of working in cardiac catheterization laboratory settings, often the responsibility for documentation, including that related to audits for quality improvement activities, is sometimes shared between the nurses performing these roles in an very inconsistent manner. Our findings show that TROOPS ratings for the same procedure were not in perfect agreement between nurses who performed these different roles. One course of action that would logically be expected to improve the reliability of measuring adverse sedation events using the TROOPS tool is to ensure that the rater is consistently only ever the nurse who performed the scout role *or* the nurse who performed the monitor role.

Few studies have evaluated the inter-rater reliability of measurement tools intended for quality improvement activities.[11] The finding that ratings between nurses using the TROOPS tool were not in perfect agreement may be a somewhat surprising finding, considering the relatively simple measurement scale. This finding underscores the need for rigorous assessment of reliability for measurement tools intended for use not only in research but also in quality improvement.

It was identified in this study that minor sedation-related adverse events related to airway and breathing were common in this population with the sedation regimen used. Gaining a baseline assessment of the frequency of such events may be useful for departments so that sedation effectiveness can be periodically evaluated when new procedures are introduced that may have different sedation requirements. The event rate is higher than most other published studies of sedation in similar cohorts.[12,13] There are no reports in the literature of the prevalence of sedation-related adverse events during procedures performed with sedation in the cardiac catheterization laboratory that have specifically used the TROOPS tool. However, serious adverse events that would meet ‘sentinel’ category definitions are rare.[14–16]

We have confirmed in this study that the time required to complete the TROOPS tool is very minimal. Using an expanded version of the TROOPS tool including items from the minor category, either routinely for all procedures, or periodically as a ‘spot-check’ for quality assurance is unlikely to interfere with clinical workflow requirements for nursing staff in the cardiac catheterization laboratory.

### Limitations

We did not collect information on which nurses completed TROOPS ratings so that anonymity was maintained. As such, the total number of nurses who completed TROOPS ratings for patients in this study is not known. For the same reason, we can not report on the characteristics of the nurses who provided ratings, such as their experience or training.[17] It should also be noted that, as expected, most events identified were from the minor category of the airway and breathing domain. Results should not be generalized outside of this context. The less common and more severe adverse events, such as the unintended requirement for tracheal intubation during sedation, is not likely to be inconsistently rated by nurses. Likewise, results should be considered most applicable to the context in which sedation was performed in this study. An approach that is common across many cardiac catheterization laboratories was used, with bolus doses of midazolam and fentanyl being administered by nurses.[18] Although there was a substantial proportion of events recorded by nurses in the sample (52%), it is possible that it over-estimates the true event rate because the analysis was performed using a convenience sample. The sample represents only a small proportion of the total number of procedures performed with sedation at the site over the data collection period. Estimating the prevalence of adverse sedation-related events was not the primary purpose of this study so we did not collect information about the total number of procedures performed with sedation over the data collection period to determine the exact proportion of completed TROOPS ratings.

### Conclusions

Paired ratings of sedation-related adverse events by nurses in the cardiac catheterization laboratory achieved moderate consistency using a version of the TROOPS tool containing items from the minor category of the airway and breathing domain. Use of this tool in clinical practice seems feasible. The time required to complete the ratings was very minimal, with most ratings taking less than one minute to complete.

Table : Participant characteristics

| Characteristic | N = 401 |
| --- | --- |
| Age | 68 (61, 76) |
| Sex |  |
| Female | 16 (40%) |
| Male | 24 (60%) |
| Fentanyl (mcg) | 100 (50, 125) |
| Midazolam (mg) | 5.0 (2.4, 8.2) |
| Procedure |  |
| Atrial fibrillation ablation | 6 (15%) |
| Coronary angiography or intervention | 4 (10%) |
| Electrophysiology study and radiofrequency ablation | 9 (22%) |
| Implantable Cardioverter Defibrillator | 7 (18%) |
| Permanent pacemaker | 8 (20%) |
| Ventricular tachycardia ablation | 6 (15%) |
| Procedure duration | 83 (36, 130) |
| BaselineOxygen (L/min) |  |
| < 6 | 6 (15%) |
| > 6 | 1 (2.5%) |
| 6 | 33 (82%) |
| 1Median (IQR); n (%) | |

# References

1. Dobson G, Chong MA, Chow L et al. Procedural sedation: A position paper of the Canadian Anesthesiologists’ Society. *Canadian Journal of Anesthesia* 2018; **65**: 1372–84.

2. Apfelbaum J, Gross J, Connis R et al. Practice guidelines for moderate procedural sedation and analgesia 2018: A report by the American Society of Anesthesiologists Task Force on Moderate Procedural Sedation and Analgesia, the American Association of Oral and Maxillofacial Surgeons, American College of Radiology, American Dental Association, American Society of Dentist Anesthesiologists, and Society of Interventional Radiology. *Anesthesiology* 2018; **128**: 437–79.

3. Werthman JA, Dietrich MS, Maxwell CA, Jordan LM, Minnick AF. The administration of moderate sedation and the role of the radiology nurse. *Journal of Radiology Nursing* 2020; **39**: 92–8.

4. Hinkelbein J, Lamperti M, Akeson J et al. European Society of Anaesthesiology and European Board of Anaesthesiology guidelines for procedural sedation and analgesia in adults. *European Journal of Anaesthesiology* 2018; **35**: 6–24.

5. Conway A, Rolley J, Page K, Fulbrook P. Clinical practice guidelines for nurse-administered procedural sedation and analgesia in the cardiac catheterization laboratory: A modified delphi study. *Journal of Advanced Nursing* 2013; **70**: 1040–53.

6. Roback M, Green S, Andolfatto G, Leroy P, Mason K. Tracking and reporting outcomes of procedural sedation (TROOPS): Standardized quality improvement and research tools from the international committee for the advancement of procedural sedation. *British Journal of Anaesthesia* 2018; **120**: 164–72.

7. Gwet KL. *Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Raters*. Advanced Analytics, LLC, 2014.

8. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2021.

9. Cravero JP, Askins N, Sriswasdi P, Tsze DS, Zurakowski D, Sinnott S. Validation of the pediatric sedation state scale. *Pediatrics* 2017; **139**.

10. White K, Macfarlane H, Hoffmann B et al. Consensus statement of standards for interventional cardiovascular nursing practice. *Heart, Lung and Circulation* 2018; **27**: 535–51.

11. Redley B, Waugh R. Mixed methods evaluation of a quality improvement and audit tool for nurse-to-nurse bedside clinical handover in ward settings. *Applied Nursing Research* 2018; **40**: 80–9.

12. Conway A, Page K, Rolley J, Fulbrook P. Risk factors for impaired respiratory function during nurse-administered procedural sedation and analgesia in the cardiac catheterisation laboratory: A matched case–control study. *European Journal of Cardiovascular Nursing* 2013; **12**: 393–9.

13. Conway A, Collins P, Chang K, Mafeld S, Sutherland J, Fingleton J. Sequence analysis of capnography waveform abnormalities during nurse-administered procedural sedation and analgesia in the cardiac catheterization laboratory. *Scientific Reports* 2019; **9**: 1–9.

14. Markham R, Martin P, Chia K, Westerink A, Denman R. Conscious sedation for electrophysiological procedures: A prospective multi centre experience. *Heart, Lung and Circulation* 2015; **24**: S226.

15. Sawhney V, Bacuetes E, Wray M et al. Moderate sedation in cardiac electrophysiology laboratory: A retrospective safety analysis. *Heart* 2017; **103**: 1210–5.

16. Conway A, Page K, Rolley JX, Worrall-Carter L. Nurse-administered procedural sedation and analgesia in the cardiac catheter laboratory: An integrative review. *International Journal of Nursing Studies* 2011; **48**: 1012–23.

17. Kottner J, Audigé L, Brorson S et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *International journal of nursing studies* 2011; **48**: 661–71.

18. Conway A, Rolley J, Page K, Fulbrook P. Trends in nurse-administered procedural sedation and analgesia across cardiac catheterisation laboratories in australia and new zealand: Results of an electronic survey. *Australian Critical Care* 2014; **27**: 4–10.