

# Direct Observation of Procedural Skills for Entry-Level Competency Assessment: A Pilot Study of Arterial Blood Gas sampling

## DOPS pour l'évaluation des compétences à l'entrée de stage: Etude pilote sur le prélèvement artériel des gaz du sang

Khedija Zaouche<sup>1</sup>, Aida Benzarti<sup>2</sup>, Fedia El Aych<sup>1</sup>, Emna Rezgui<sup>1</sup>, Emna Kallel<sup>1</sup>, Manel Kallel<sup>1</sup>

1. Emergency Department, Mahmoud El Matri Hospital, Faculty of Medicine of Tunis, University of Tunis El Manar, Tunis, Tunisia

2. Department of Occupational Medicine, la Rabta Hospital, Pedagogical Committee, Faculty of Medicine of Tunis, University of Tunis El Manar.

### ABSTRACT

**Introduction:** Proficiency in arterial blood gas (ABG) sampling is expected of senior medical students, yet it is often assumed rather than objectively assessed. Early identification of procedural skill gaps is essential to ensure patient safety, particularly in emergency settings. This study aimed to evaluate the value of Direct Observation of Procedural Skills (DOPS) as a diagnostic tool at the start of an emergency medicine clerkship.

**Methods:** In this prospective pilot study, 21 fifth-year medical students were assessed while performing ABG sampling on real patients using a standardized DOPS checklist. The primary outcome was the level of supervision required. Secondary outcomes included performance in technical and non-technical domains, prior procedural experience, and self-reported stress.

**Results:** Sixteen students (76%) required direct supervision, while only one (5%) achieved near-independent performance. Technical skills were generally satisfactory; however, major deficiencies were observed in informed consent (71%) and communication (62%). Prior experience of  $\geq 10$  procedures was significantly associated with higher competence ( $p < 0.01$ ), whereas no association was found with stress levels.

**Conclusion:** Procedural competence in ABG sampling is heterogeneous among senior students. Early implementation of DOPS helps identify gaps, particularly in non-technical domains, and supports tailored supervision to enhance patient safety.

**Keywords:** Clinical Competence; Undergraduate Medical Education; Feedback; Emergency Medicine.

### RÉSUMÉ

**Introduction:** La maîtrise du prélèvement des gaz du sang artériel (GDS) est attendue chez les étudiants en fin de cursus, mais elle demeure souvent supposée plutôt qu'objectivement évaluée. L'identification précoce des lacunes procédurales est essentielle pour garantir la sécurité des patients, notamment aux urgences. Cette étude visait à évaluer l'intérêt du Direct Observation of Procedural Skills (DOPS) comme outil diagnostique à l'entrée d'un stage de médecine d'urgence.

**Méthodes:** Dans cette étude pilote prospective, 21 étudiants de cinquième année ont été évalués lors d'un prélèvement de GDS sur patients réels à l'aide d'une grille DOPS standardisée. Le critère principal était le niveau de supervision requis. Les critères secondaires incluaient la performance dans les domaines techniques et non techniques, l'expérience préalable et le stress auto-rapporté.

**Résultats:** Seize étudiants (76 %) nécessitaient une supervision directe, tandis qu'un seul (5 %) atteignait un niveau quasi autonome. Les compétences techniques étaient globalement satisfaisantes, mais des déficits majeurs concernaient le consentement éclairé (71 %) et la communication (62 %). Une expérience  $\geq 10$  procédures était associée à une meilleure compétence ( $p < 0,01$ ), sans lien avec le stress.

**Conclusion:** La compétence en GDS est hétérogène chez les étudiants seniors. Le DOPS précoce permet d'identifier les lacunes, notamment non techniques, et d'adapter la supervision pour renforcer la sécurité des patients.

**Mots-clés:** Compétence clinique ; Formation médicale de premier cycle ; Feedback ; Médecine d'urgence.

### Correspondance

Khedija Zaouche

Emergency Department, Mahmoud El Matri Hospital, Faculty of Medicine of Tunis, University of Tunis El Manar, Tunis, Tunisia

Email: khedija.zaouche@fmt.utm.tn

## INTRODUCTION

Ensuring that medical students are competent to perform invasive procedures is a core responsibility of clinical educators, particularly in high-stakes environments like the Emergency Department (ED). While competency-based education has shifted the focus toward workplace-based assessment [1], the transition from "knowing how" to "doing" remains a critical period for patient safety. Arterial blood gas (ABG) sampling is a fundamental skill expected of senior students. It requires technical accuracy, strict adherence to asepsis, and effective communication, yet it is often painful and carries risks of complications [2]. Traditionally, supervision in busy EDs can be variable, with students often performing procedures based on self-reported confidence rather than demonstrated competence. Direct Observation of Procedural Skills (DOPS) is a well-established tool for assessing procedural skills on real patients [3–5]. However, while DOPS is frequently used for summative assessment at the end of rotations, its potential as a baseline diagnostic tool upon entry to a clerkship is less explored, particularly in the context of emergency medicine. Identifying learning needs before students engage in independent practice is crucial to mitigate learner stress and ensure patient safety [6, 7]. This study aimed to use DOPS to assess the baseline procedural competence of medical students in arterial blood gas sampling at the start of their emergency medicine clerkship. The objective was to uncover specific skill gaps to inform tailored supervision and feedback strategies.

## METHODS

### Study Design and Setting

This prospective descriptive pilot study was conducted in the Emergency Department of a university-affiliated teaching hospital over a six-week period (August–September 2025). The study adheres to the STROBE guidelines for observational studies.

### Participants

Participants were fifth-year medical students undertaking their emergency medicine clerkship. Inclusion criteria were students who had received prior theoretical instruction and had observed the procedure at least once. We excluded preclinical students and those who had never attempted the procedure, as the focus was on assessing baseline competence in a clinical setting. Written informed consent was obtained from all participating students and patients.

### The Instrument (DOPS)

We used a standardized DOPS checklist validated in the literature (Supplementary file 1). The tool assessed ten domains, including knowledge (indications/anatomy), informed consent and communication,

equipment preparation, aseptic technique and safety, technical ability, and post-procedure management and documentation.

### Data Collection

Procedure assessments were conducted in a dedicated procedure room on real patients requiring ABG sampling as part of standard care. Before radial artery puncture, students were expected to assess the indication, check for contraindications, identify the puncture site, and verify collateral circulation using the Allen test when appropriate. To minimize assessor subjectivity, two academic emergency physicians with formal training in simulation and debriefing observed each encounter simultaneously. The assessment followed a strict three-step sequence: (a) Briefing (3 min): Explanation of the assessment process to the student, (b) Observation (5 min): The student performed the procedure (patient preparation, asepsis, puncture, labeling) without interruption and (c) Debriefing (10 min): Immediate structured feedback was provided, focusing on technical skills, communication, and safety.

### Outcome Measures

The primary outcome was the level of supervision required. Performance was graded on a standard ordinal scale, with the target defined as Level 4b: "Independent and confident execution with anticipation of common problems". The secondary outcome was self-reported stress, measured immediately post-procedure using a generic Numerical Rating Scale (0–10).

### Statistical Analysis

Data were analyzed using SPSS version 21. Categorical variables are presented as frequencies and percentages, and continuous variables as medians with interquartile ranges (IQR). Given the small sample size and ordinal nature of the data (DOPS levels, stress scores), the Mann–Whitney U test was used to compare groups based on prior experience (<10 vs. ≥10 procedures). A p-value <0.05 was considered statistically significant.

### Ethical Approval

The study protocol was approved by the local ethics committee (Ref: 2/2025). Informed consent was obtained from all participants involved in the study.

## RESULTS

### Participant Characteristics

A total of 21 fifth-year medical students participated in the study. Prior exposure to arterial blood gas (ABG) sampling was limited and heterogeneous. Nearly half of the participants (43%, n=9) performed the procedure only once prior to the assessment. A minority of students

(24%, n=5) reported more extensive experience, having performed the procedure ten times or more [Table 1].

**Table 1.** Participant characteristics and prior procedural experience

| Participant characteristics                                | Value      |
|------------------------------------------------------------|------------|
| Age, mean $\pm$ SD (years)                                 | 22 $\pm$ 2 |
| <b>Sex, n (%)</b>                                          |            |
| • Male                                                     | 7 (33)     |
| • Female                                                   | 14 (67)    |
| <b>Prior arterial blood gas sampling experience, n (%)</b> |            |
| • 1 prior attempt                                          | 9 (43)     |
| • 2–5 prior attempts                                       | 7 (33)     |
| • $\geq 10$ prior attempts                                 | 5 (24)     |
| <b>Prior exposure during training*, n (%)</b>              |            |
| • Simulation-based session                                 | 2 (10)     |
| • Supervised demonstration                                 | 7 (33)     |

\*Some students reported more than one type of prior exposure; n: number; % = percentage within each group.

### Baseline Procedural Competence

Overall, procedural autonomy was limited at the start of the clerkship, with 16 students (76%) classified at novice levels and requiring direct supervision to ensure patient safety. Specifically, 6 students were rated at Level 2a and 10 at Level 2b. Intermediate competence was observed in only 4 participants (Level 3a or 3b), and only a single student (5%) demonstrated near-independent execution (Level 4a). No student achieved the highest level of mastery (Level 4b).

### Domain Analysis: Technical vs. Non-Technical Skills

The DOPS assessment revealed a dissociation between technical proficiency and non-technical skills [Table 2].

**Table 2.** Baseline procedural performance by domain

| Domain assessed                                 | Development required n (%) | Satisfactory n (%) | Excellent n (%) |
|-------------------------------------------------|----------------------------|--------------------|-----------------|
| Knowledge (indications, anatomy, complications) | 11 (52)                    | 10 (48)            | 0               |
| Informed consent                                | 15 (71)                    | 6 (29)             | 0               |
| Equipment preparation                           | 4 (19)                     | 16 (76)            | 1 (5)           |
| Asepsis and safety                              | 8 (38)                     | 10 (48)            | 3 (14)          |
| Technical execution                             | 9 (43)                     | 11 (52)            | 1 (5)           |
| Management of unexpected events                 | 11 (52)                    | 9 (43)             | 1 (5)           |
| Documentation                                   | 9 (43)                     | 11 (52)            | 1 (5)           |
| Communication                                   | 13 (62)                    | 7 (33)             | 1 (5)           |
| Professional behavior                           | 8 (38)                     | 12 (57)            | 1 (5)           |

n : number, % = percentage within each group.

Students performed relatively well in tangible procedural tasks, particularly in equipment preparation, where 76% were rated satisfactory or excellent, and in aseptic technique, with 62% achieving satisfactory performance.

However, substantial deficiencies were observed in non-technical domains. A majority of students (71%) required further development in obtaining informed consent, and 62% demonstrated inadequate communication with the patient during the procedure, highlighting important gaps in patient-centered competencies despite acceptable technical execution.

### Management of unexpected events

Regarding the management of unexpected events, 52% of students were unprepared to handle complications. Prior clinical experience was a significant predictor of competence. Students who had performed the procedure at least ten times consistently achieved higher performance levels (clustering in Levels 3 and 4), whereas those with fewer than 10 attempts remained predominantly at novice levels (Levels 2a and 2b) ( $p = 0.005$ ).

In contrast, self-reported stress levels immediately post-procedure (median 6/10) did not correlate with performance quality. Elevated stress scores (range 6–8) were prevalent across all proficiency categories, with novice students (Level 2) reporting anxiety levels comparable to those demonstrating higher independence.

## DISCUSSION

### The "Competence Gap" in Senior Students

Our findings highlight a critical disconnect between academic seniority and procedural autonomy. Despite being in their fifth year of medical school, most students demonstrated limited baseline competence. This heterogeneity reflects variable prior clinical exposure rather than theoretical knowledge. Crucially, the traditional "see one, do one" model appears insufficient for complex invasive procedures; our data suggest students with  $\geq 10$  prior procedures performed better; however, this threshold should be interpreted cautiously given the small sample size and requires confirmation in larger studies [8, 9].

### Dissociation of Skills: The "Blind Spots"

A key finding of this study was the clear dissociation between technical and non-technical performance. While students could often prepare equipment and perform the puncture, they frequently neglected essential non-technical domains such as informed consent, communication, and safety checks. This aligns with previous literature suggesting that novices tend to focus their cognitive load on the manual task at the expense of patient interaction [3, 10]. Such variability in performance across different assessment domains is consistent with previous reviews of DOPS validity [11]. Identifying these deficits is crucial, as literature suggests that specific interventions, such as simulation training to increase error awareness [12] and reflective portfolios

to track competence [13], are often required to address such non-technical gaps effectively. Using DOPS early in the rotation effectively flagged these specific "blind spots," which might have otherwise gone unnoticed in a busy emergency department focused on successful sampling rather than holistic care.

### Stress and Assessment

Although students reported moderate to high stress levels (median 6/10), this anxiety did not correlate with poor performance. This supports the view that while workplace-based assessment induces anxiety, it does not necessarily impair competence if the assessment is framed correctly as a learning opportunity [6, 14]. While excessive stress can negatively affect performance [15], stress in this context may simply reflect the high-stakes nature of performing invasive procedures on real patients rather than the burden of the assessment tool itself.

### Limitations

This study has several limitations. First, it was a single-center pilot study with a small sample size (n=21), which limits the generalizability of the statistical findings. Second, the lack of inter-rater reliability testing or formal assessor calibration and the absence of longitudinal follow-up prevent us from drawing conclusions about the long-term retention of skills or the precise consistency of scoring. Third, patient-level anatomical or vascular factors potentially influencing puncture difficulty, such as radial artery diameter or ultrasound-based assessment, were not collected, as the study focused primarily on students' procedural competence rather than determinants of technical success. However, as a baseline "diagnostic" snapshot, the sample was sufficient to reveal prevailing training needs and justify the intervention.

### Contribution to the literature

This study highlights that academic seniority does not guarantee procedural competence in invasive skills like arterial blood gas sampling. It demonstrates that using Direct Observation of Procedural Skills (DOPS) as a baseline diagnostic tool effectively identifies specific deficits, particularly in non-technical domains such as communication and informed consent. These findings support shifting from generic supervision to targeted guidance based on demonstrated needs to ensure patient safety.

### Implications for Practice

For clinical educators, this study confirms that procedural competence should not be assumed based on a student's year of study. Implementing DOPS at the start of a clerkship—rather than solely as a final summative exam—transforms it into a powerful diagnostic tool. It allows supervisors to objectively identify "unsafe" novices who require direct supervision, regardless of their self-reported confidence, and to tailor feedback

toward neglected non-technical domains.

## CONCLUSION

Implementing DOPS at the entry of an emergency medicine clerkship proved to be a valuable diagnostic strategy. It revealed that procedural competence in senior students is highly variable. By objectively identifying gaps—particularly in non-technical domains—early DOPS assessment allows educators to move from generic supervision to targeted guidance based on actual demonstrated needs, ensuring patient safety during the critical transition to independent practice.

## REFERENCES

1. Monga S, Agarwal D. Evaluation of direct observation of procedural skills (DOPS) as an assessment method for MBBS students. *Res J Med Sci.* 2024;18:436–440. <https://doi.org/10.36478/makrjms.2024.12.436.440>
2. Cunanan B, Muppa H, Orellana MA, Bates RN, McGain F. Blood gas sampling in the intensive care unit: a prospective before-and-after interventional study on the effect of an educational program on blood gas testing frequency. *Aust Crit Care.* 2024;37(5):755–760. <https://doi.org/10.1016/j.aucc.2024.01.009>
3. Erfani Khanghahi M, Ebadi Fard Azar F. Direct observation of procedural skills (DOPS) evaluation method: a systematic review of evidence. *Med J Islam Repub Iran.* 2018;32:45. <https://doi.org/10.14196/mjiri.32.45>
4. Asadi K, Mirbolook AR, Haghighi M, Sedighinejad N, Naderi Nabi B, Abedi S, et al. Evaluation of satisfaction level of orthopedic interns from direct observation of procedural skills assessment (DOPS). *Res Med Educ.* 2012;4(2):17–23.
5. Watson MJ, Wong DM, Kluger R, Chuan A, Herrick MD, Ng I, et al. Psychometric evaluation of a direct observation of procedural skills assessment tool for ultrasound-guided regional anaesthesia. *Anaesthesia.* 2014;69(6):604–612. <https://doi.org/10.1111/anae.12625>
6. Cohen SN, Farrant PB, Taibjee SM. Assessing the assessments: UK dermatology trainees' views of the workplace assessment tools. *Br J Dermatol.* 2009;161(1):34–39. <https://doi.org/10.1111/j.1365-2133.2009.09097.x>
7. Bindal N, Goodyear H, Bindal T, Wall D. Direct observation of procedural skills (DOPS) assessment: a study to evaluate the experience and opinions of trainees and assessors. *Med Teach.* 2013;35(6):e1230–e1234. <https://doi.org/10.3109/0142159X.2012.746447>
8. Dabaghi Tabriz F, Bahramian A, Asdaghi S, Ghadimi F, Katebi K. The effect of direct observation of procedural skills on learning and satisfaction of students in a community dentistry course. *J Adv Med Educ Prof.* 2024;12(2):111–117. <https://doi.org/10.30476/JAMP.2023.100431.1890>
9. Lagoo JY, Joshi SB. Introduction of direct observation of procedural skills (DOPS) as a formative assessment tool during postgraduate training in anaesthesiology: exploration of perceptions. *Indian J Anaesth.* 2021;65:202–209. [https://doi.org/10.4103/ija.ija\\_124\\_20](https://doi.org/10.4103/ija.ija_124_20)
10. Choudhary A, Gupta V. Teaching communications skills to medical students: introducing the fine art of medical practice. *Int J Appl Basic Med Res.* 2015;5(Suppl 1):S41–S44. <https://doi.org/10.4103/2229-516X.162273>
11. Rela M, Price T. Review of the validity of direct observation of procedural skills (DOPS) as an assessment tool for the procedural skills of surgical trainees. *Ann R Coll Surg Engl.* 2023;105(7):599–606. <https://doi.org/10.1308/rcsann.2022.0052>
12. Yalcinturk AA, Ustun B. Medical errors room: the effect of simulation training on increasing awareness of nursing students about medical

- errors encountered in psychiatric clinics. *Arch Psychiatr Nurs*. 2023;43:37–42. <https://doi.org/10.1016/j.apnu.2022.12.030>
13. Haldane T. "Portfolios" as a method of assessment in medical education. *Gastroenterol Hepatol Bed Bench*. 2014;7(2):89-93.
  14. Cobb K, Brown G, Jaarsma D, Hammond R. The educational impact of assessment: a comparison of direct observation of procedural skills and multiple-choice questions. *Med Teach*. 2013;35(11): e1598–e1607.
  15. <https://doi.org/10.3109/0142159X.2013.803061>
  16. Amini A, Shirzad F, Mohseni MA, Sadeghpour A, Elmi A. Designing direct observation of procedural skills (DOPS) test for selective skills of orthopedic residents and evaluating its effects from their points of view. *Res Dev Med Educ*. 2015;4(2):147-152. <https://doi.org/10.15171/rdme.2015.026>