

Evaluating Pain in Intensive Care

Hanna SUOMINEN^a, Heljä LUNDGRÉN-LAINE^b,
Sanna SALANTERÄ^b and Tapio SALAKOSKI^a

^a*Turku Centre for Computer Science (TUCS) and Department of Information
Technology, University of Turku, Turku, Finland*

^b*Department of Nursing Science, University of Turku, Turku, Finland*

Abstract. Optimal pain management is essential for good care outcomes, but assessing pain is particularly complex in intensive care, as patients are often unable to communicate. We hypothesize that the task could be supported through human language technology. To evaluate the feasibility of such tools, we study how pain is documented in electronic Finnish free-text intensive care nursing notes by statistically comparing annotations of ten nursing professionals on a set of 1548 documents. The aspects considered include the amount and writing style of pain-related notes, pain intensity, and given pain care. More than half of the documents contained information relevant for patients' pain status but it was expressed usually indirectly. Also pain medication was commented as free-text. Although annotators' pain intensity evaluations diverged, the substantial amount of pain-related notes encourages developing computational tools for pain assessment.

Keywords: computerized patient records, decision-making, intensive care, natural language processing, pain

1. Introduction

Optimal pain management in intensive care (IC) is critical but difficult: Imperfect assessment and treatment of pain is found to be associated to severe physiological consequences (e.g., myocardial ischemia, atelectasis, pneumonia) and psychological problems such as post-traumatic stress [1], [2]. Through systematic pain assessment, incidences of pain and nosocomial infections as well as duration of mechanical ventilation can be decreased [3]. However, IC patients' inability to communicate often forces nurses to perform the evaluation only relying on implicit physiological and behavioural indicators without a possibility for consulting the patient [4]. In addition, there is evidence on IC patients' dissatisfaction with their pain control [5], [6].

We hypothesize that pain assessment can be supported through human language technology, and in this paper, we evaluate the feasibility of such tools: We study how pain is documented in free-text IC nursing notes. The aspects considered include amount, writing style, pain intensity, and given pain care. If the notes contain data valuable for evaluating pain (e.g., signals of pain and pain medication efficiency), a second opinion for the patient's pain status could be generated automatically.

2. Material and Methods

The data consisted of nursing notes of 516 adult long-term IC patients. These electronic Finnish records covered the whole in-patient time and they were written mainly for intra-unit information exchange. The whole data set included 17140 patient and shift-specific documents (Figure 1) with an average of 73 tokens (including punctuation) per document. We chose long-term patients (i.e., length of stay at least five days), as use of

Consciousness: Conscious and oriented in the evening. Fuzzy already at night-time but luckily not very nervous. Fell asleep with SOPORIFIC_DRUG_NAME.

Hemodynamics: Pressure level low at all times and filled plenty regardless of anuria. Because of the fasciotomy wound needs constantly blood and fluids in the small hours to maintain the pressure level. Heart rate level varies according to the pressure level, sr.

Rhabdomyolysis: Dialysis performed in the evening which removed 1000 ml. There has been no own diuresis. In the evening got DIURETIC_DRUG_NAME boluses but they had no effect. Left leg at midnight rock-solid from the calf and toes were not moving anymore. A surgeon visited and the patient to the operating theatre for the fasciotomy. Fasciotomies at both sides of the calf and one cut in the thigh. Wounds bled a lot why constant replacement with erythrocytes and in addition fluids.

Good oxidation with ventilation mask.

Figure 1. An anonymous, typographical error preserving example translated from Finnish to English.

their nursing notes to support pain evaluation is likely to be hindered due to the large quantity of text. Before collecting the de-identified data retrospectively from 2005 Jan 1 to 2006 Aug 1, the proper permissions¹ were furnished for the study protocol.

Ten annotators analyzed independently the content of 1548 documents; 3 documents per patient were chosen randomly. Annotators had varying education, nursing experience and pain expertise² (Table 1). The analysis included for every document

1. **Directness of expressions**
 - 0 – Pain/painlessness is not described at all 1 – It is described only indirectly
 - 2 – It is described only directly 3 – It is described both indirectly and directly
2. **Patient's pain intensity based on a ten point numerical rating scale³**
 - 0 – No pain (i.e., score 0) 1 – Mild pain (i.e., score 1–3)
 - 2 – Moderate pain (i.e., pain score 4–5) 3 – Severe pain (i.e., pain score 6–7)
 - 4 – Very severe or worst possible pain (i.e., pain score 8–10)
3. **Personnel's response⁴**
 - 0 – Not described
 - 1 – Pain medication or pain medication combined with other pain relief methods
 - 2 – Other pain relief methods without pain medication
 - 3 – Neither pain medication nor pain medication combined with other pain relief methods.

For the first two items, also confidence in the judgement was evaluated by using dichotomy between certain and uncertain answers.

Written instructions and the first author guided the annotation process. Before the task, annotators and authors practiced together with an exercise set of ten documents. Annotators were encouraged to use their expertise, when interpreting the notes and, for example, reference books for medical trademarks were allowed. However, using other patient specific information than that provided in the document in question was forbidden. On average, 19 hours and 20 minutes was used to complete the actual annotation task (Table 1).

Distribution of data to different evaluation aspects and their subcategories was studied through frequencies. Inter-annotator agreement in the Directness of expressions and the Personnel's response was measured by using Cohen's κ [7] for pair-wise comparisons

¹ Statutes of Finland: Medical research act 488/1999 and decree 986/1999.

² Evaluated based on pain education, pain studying and involvement in special pain evaluation groups during the last two years.

³ Evaluated only, when the previous answer was 1–3. The scale was chosen as it is one of the most used tools for pain intensity assessment.

⁴ Evaluated only, when the previous answer was 1–4.

Table 1. Descriptive statistics of annotators.

ID	Education	Nursing experience	Pain expertise ⁵	Annotation hours
1	RN	6 years in adults' IC	high	18
2	RN	12 years in adults' IC	high	11
3	RN	17 years in adults' IC	high	25
4	RN	18 years in adults' IC	high	30
5	RN, 1 st year MNSc student	13 years in adults' IC	intermediate	20
6	RN, 1 st year MNSc student	16 years in adults' IC	intermediate	13
7	RN, 3 rd year MNSc student	21 years in surgical wards	intermediate	20
8	RN, 3 rd year MNSc student	1 year in adults' IC, 12 years in paediatrics	intermediate	20
9	RN, MNSc, PhD student	16 years in adults' IC	intermediate	16
10	RN, MNSc, PhD student	7 years in adults' IC	intermediate	20

and Fleiss' κ [8] for groups of more than two annotators. In group-wise comparisons, education (RN vs. academic studies), nursing experience (five- and ten-year intervals) and pain expertise (intermediate or high) were selected as distinguishing features. With these statistics, $\kappa < 0$ for poor agreement and its values between 0 and 1 denote agreement. Kendall's τ_b [9, pp. 40–42], ranging between -1 for perfect disagreement and 1 for perfect agreement, was used to measure rank correlation in pain intensity assessments. For multiple hypothesis testing, a significance level was Bonferroni corrected to $\alpha = 0.005/45 \approx 0.0002$. To eliminate accumulation of previous disagreements, only the documents that all nurses had evaluated were considered for the latter evaluation aspects. R 2.6.2 was used for all analyses.

3. Results

In annotations, 51% of the documents described patients' pain status; the most substantial proportion being indirect statements (Figure 2, 36%). On average 17% of answers were uncertain, minimum, maximum and standard deviation being 2% (Table 1, ID 4), 66% (ID 3) and 19%, respectively. Based on confidence scores, evaluating if pain or painlessness is described indirectly seemed to be the most challenging task.

Agreement on the Directness of expressions was fair (Fleiss' κ 0.25), and category-wise, annotators better agreed on cases, where pain was described only directly or not at all (Fleiss' κ 0.37 and 0.29, respectively) than, when an indirect pain description was present (Fleiss' κ 0.16 and 0.12 for the answer options 1 and 3, respectively). Considering differences in education, nursing experience and pain expertise did not substantially improve the agreement; instead of professional characteristics, the task was likely dominated by individual variation in language interpretation. When focusing on pair-wise agreement of different annotations, the average was only 0.27, the minimum being almost zero and maximum moderate 0.50 (Table 2).

A clear majority of cases were evaluated as the patient having moderate or severe pain (Figure 2). Pain intensity evaluation was very difficult for the annotators; the average uncertainty proportion was as large as 77% with the standard deviation of 24%, minimum of 32% (Table 1, ID 8) and maximum of 98% (ID 3). Different pain intensity annotations correlated positively but with a moderate degree, Kendall's τ_b being commonly about 0.50 with the minimum of 0.28 and maximum of 0.64 (Table 2).

⁵ Scale: low–intermediate–high.

4. Conclusion

Our IC nursing documentation analysis suggests that more than half of the notes contain information relevant for patients' pain status and also pain medication is described as free-text. Evidence supporting the claim has also been reported: [10] studies records of 52 critically ill Canadian patients to describe pain indicators and management. It provides evidence of nursing notes addressing pain medication and behavioural pain indicators, and physicians documenting seldom any pain assessment. Our previous study [11] considers nursing notes of 43 Finnish IC patients and concludes that they contain expressions relevant to pain evaluation. Compared to [10] and [11], this study, based on opinions of ten nursing professionals, reflects more reliably current pain documentation practices in Finnish electronic IC nursing notes.

Using information documented as free-text to support decision-making in pain care is likely to be challenging: First, the substantial amount of notes can make extraction of the relevant parts difficult. Second, pain is typically expressed indirectly. And third, the annotators have relatively differing opinions on pain intensity. However, through learning from all annotations, a computational tool may be able to portray cumulative pain knowledge mined from the individual opinions and form a more standard reference for pain intensity assessment than consulting another nurse would.

By building on this study, our future aim is to develop human language technology for pain assessment. The output could serve as one evaluation of the current pain status and hence confirm the personnel's view. Further, it could be used for generating reminders, trends describing time-wise pain status, and alarms related to pain management.

Acknowledgements

We express our gratitude to the Finnish Funding Agency for Technology and Innovation, Tekes (grant 40020/07) as well as gatherers and annotators of the data.

References

- [1] McArdle P. Intravenous analgesia. *Crit Care Clin* 1999; 15: 89–104.
- [2] Rotondi A, Chelluri L, Sirio C, Mendelsohn A, Schulz R, Belle S, Im K, Donahoe M, Pinsky MR. Patients' recollections of stressful experiences while receiving prolonged mechanical ventilation in an intensive care unit. *Crit Care Med* 2002; 30: 746–752.
- [3] Chanques G, Jaber S, Barbotte E, Violet S, Sebbane M, Perrigault P-F, Mann C, Lefrant J-Y, Eledjam J-J. Impact of systematic evaluation of pain and agitation in an intensive care unit. *Crit Care Med* 2006; 34: 1691–1699.
- [4] Sessler CN, Grap MJ, Ramsay MAE. Evaluating and monitoring analgesia and sedation in the intensive care unit. *Crit Care* 2008; 12: S2.
- [5] Hamill-Ruth RJ, Marohn ML. Evaluation of pain in the critically ill patient. *Crit Care Clin* 1999; 15: 35–54, v–vi.
- [6] Puntillo KA, Wild LR, Morris AB, Stanik-Hutt J, Thompson CL, White C. Practices and predictors of analgesic interventions for adults undergoing painful procedures. *Am J Crit Care* 2002; 11: 415–429, quiz 430–431.
- [7] Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960; 20: 37–46.
- [8] Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971; 76: 378–382.
- [9] Kendall M, Gibbons JD. *Rank Correlation Methods* (5th ed.). London, UK: Edward Arnold; 1990.
- [10] Gélinas C, Frontier M, Viens C, Fillion L, Puntillo K. Pain assessment in critically ill intubated patients: a retrospective study. *Am J Crit Care* 2004; 13: 126–135.
- [11] Suominen H, Pahikkala T, Hiisa M, Lehtikunnas T, Back B, Karsten H, Salanterä S, Salakoski T. Relevance ranking of intensive care nursing narratives. *Lecture Notes in Computer Science* 2006; 4251: 720–727.

Email address for correspondence hajasu@utu.fi