Identifying Regional Anesthesia Procedures in the EHR Using Natural Language Processing

Laura A. Graham, PhD, MPH

U.S. Department of Veterans Affairs

Regional Anesthesia & Pain Medicine

Brief technical report

Use of natural language processing method to identify regional anesthesia from clinical notes

Laura A Graham [1], Samantha S Illarmo [2], Sherry M Wren [3,4], Michelle C Odden [2,5] and Seshadri C Mudumbai [6,7]

Correspondence to Dr Laura A Graham, Health Economics Resource Center, VA Palo Alto Health Care System, Menlo Park, California, USA; lagraham@stanford.edu

Abstract

Introduction Accurate data capture is integral for research and quality improvement efforts.
Unfortunately, limited guidance for defining and documenting regional anesthesia has resulted in wide

Overview

- Overview of NLP Approaches
- Developing the Framework and Algorithm
- Our Findings
- Summary

Poll

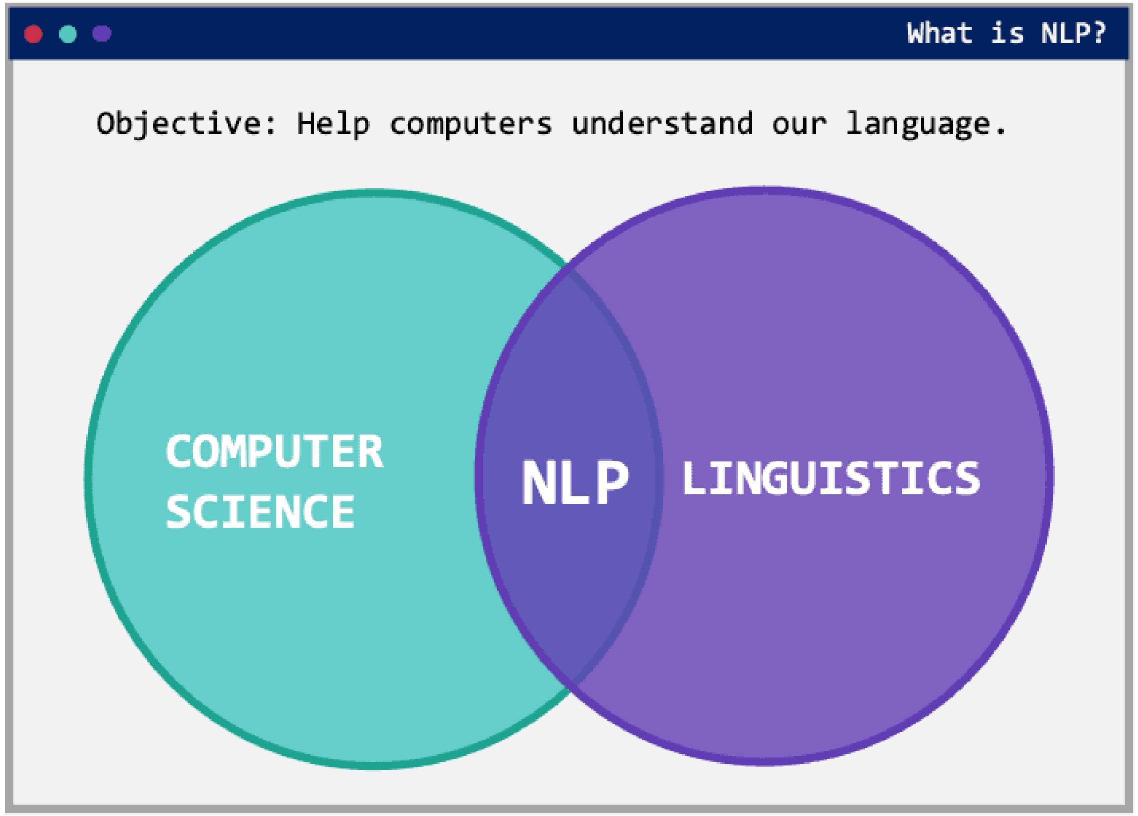
- How much experience do you have with natural language processing (NLP) methods?
  - No experience
  - A little experience
  - A decent amount of experience, but I’m no expert
  - I’m an expert

Artificial Intelligence

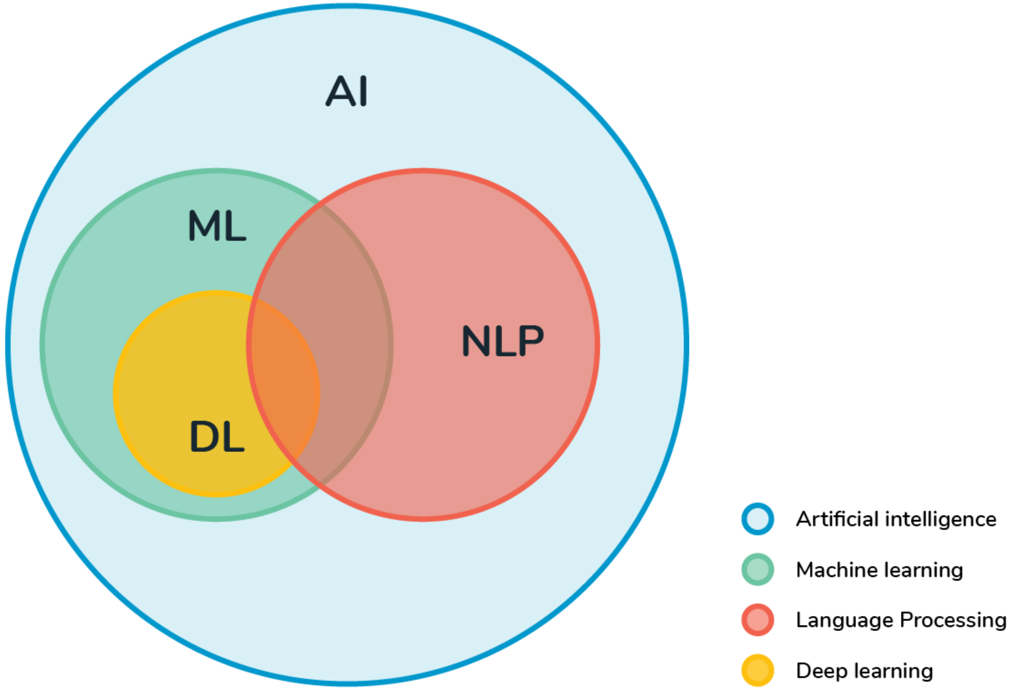
- Artificial intelligence (AI) is the **simulation of human intelligence** processes by machines, especially computer systems.
  – Broad concept of creating intelligent machines that can perform tasks that would typically require human intelligence, such as reasoning, learning, problem-solving, perception, and decision-making.
- Natural language processing (NLP) is a **subfield of AI** that focuses on enabling machines to understand, interpret, and generate human language.
  – NLP leverages AI techniques and methodologies to process and analyze natural language data.

NLP in Health Services Research

- NLP involves using algorithms to analyze, understand, and **derive meaning from human language**, most commonly used in the electronic health record (EHR).
- NLP plays a crucial role in unlocking the potential of EHRs by enabling the extraction of relevant information from unstructured textual data like clinic visit summaries or procedure descriptions in operative notes.



Approaches to NLP



Rule-based Language Processing

Machine-learning

Hybrid

Rule-Based NLP

- Rule-based natural language processing (NLP) is one of the **earliest approaches** to NLP.
- Involve developing a **set of rules or patterns** that capture specific language structures, syntax, semantics, or other linguistic phenomena.
- Rules are then applied to the input text data to extract information, classify text, or perform other NLP tasks based on the matched patterns.
- Typically developed by or with the help of domain specialists who have in-depth knowledge of the language and the problem domain.

Rule-Based NLP

- Useful when there is **limited annotated data** available for training machine learning models.
- It can provide **high precision for narrow, well-defined tasks** within specific domains where the rules can be carefully crafted.
- They can be **computationally efficient** and provide fast processing times, making them suitable for applications with strict performance requirements.
- They **require significant effort** in manually crafting and maintaining the rules, which can be time-consuming and challenging, especially for complex language phenomena.

Machine-Learning NLP

- Machine learning models learn patterns and relationships from **large annotated datasets** rather than relying on predefined rules.
  - Statistical models, neural networks, and deep learning approaches
- Can handle **complex language phenomena** and generalize by learning from examples.
- However, they require large amounts of **accurate training data**, which can be expensive and time-consuming to develop.
- The models can be opaque and lack interpretability, which makes it difficult to understand the reasoning behind their results.
- **Computationally intensive**, especially for deep learning models, requiring significant hardware resources.

Summary of NLP Approaches

- In practice, many NLP systems employ a hybrid approach, **combining rule-based components for specific subtasks with machine learning models** for more complex aspects.
- Rule-based methods can provide initial structure or constraints, while machine learning handles the generalization and adaptation to real-world data variations.
- The choice between rule-based and machine-learning approaches depends on the specific NLP task, the availability of annotated data, the required level of interpretability, and the computational resources at hand.
- In general, machine learning techniques have become more prevalent due to their ability to handle complexity and generalize.
- However, **rule-based methods can still be valuable in certain domains or as complementary components**.

Developing the Framework and Algorithm

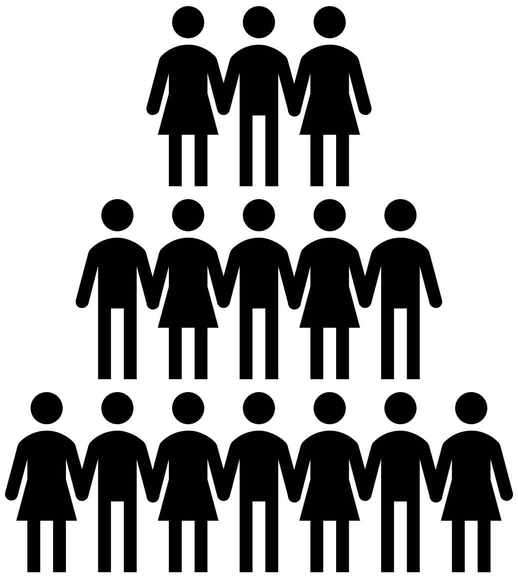
U.S. Department of Veterans Affairs

Study Population

- Patients undergoing **elective non-cardiac surgery** with general anesthesia

- January 1, 2017, and December 31, 2022

- 6 VA hospitals in VISN 21



Regional Anesthesia

- Regional anesthesia (RA) involves the administration of **anesthetic agents to a specific part of the body.**
  - The goal is to block nerve impulses and provide pain relief during and after surgery.
- This includes **spinal**, **epidural**, and nerve **blocks**, each targeting different areas of the body to achieve localized anesthesia and pain management.
- Regional techniques offer reduced risk of complications, improved postoperative pain control, and faster recovery compared to general anesthesia.

Background

- **RA is part of multimodal analgesia (MMA)**,
  - Form of pain management that includes non-opioid medications

- The goal of our larger study is to examine the effect of MMA on postoperative pain and opioid use.

- So, we **need to identify RA procedures**.

The Problem

- We initially proposed using **procedure codes** to identify RA procedures.
- There is a lot of variability in how surgical anesthesia is documented depending on when and what procedures happen.
  - We were interested in only the pre- and intraoperative period, not the postoperative period, and intraoperative anesthesia procedures are not billed by CPT code.
- Ultimately, only 1.5% of all cases had an anesthesia code.
  - We expected 25-30% for our cohort.

https://www.asra.com/news-publications/asra-newsletter/newsletter-item/asra-news/2020/02/07/regional-anesthesia-billing-surgical-anesthesia-versus-postoperative-analgesia
Cozowicz, C., J. Poeran, and S. G. Memtsoudis. "Epidemiology, trends, and disparities in regional anaesthesia for orthopaedic surgery." BJA: British Journal of Anaesthesia 115.suppl_2 (2015): ii57-ii67.

The Problem

- So, we pivoted to a variable for the type of anesthesia captured in the **CDW Surgery Domain**.
  - G (general), R (regional), M (monitored anesthesia care), E (epidural), S (spinal), and L (local)
    - Note: nerve blocks were not documented in the CDW either.
- This identified more RA procedures (16%), but the numbers still seemed low to our clinical experts on the study.

Anesthesia Data in the VA

- So, we went **back to the source** – the EHR – and began reviewing documentation around the time of surgery.
  - What types of notes were used?
  - Where was anesthesia being recorded?
  - How did the documentation overlap with our current data?
  - Did it look like there was consistent documentation across all sites, years, types of procedures …?

Data Structure

Patient was taken to the operating room and placed on the operating room table in the supine position. He had previously undergone a peripheral nerve block in the postoperative area. He then underwent a spinal anesthetic.

Anesthesia was established (Spinal Anesthesia and MAC) and the extremity was prepped using chlorhexidine and chloraprep and draped in the usual sterile fashion. A surgical time out …

… injected into all of the wounds, and the patient received a tap block at the very beginning of the case as well.

ANESTHETIC: General + regional. thoracic epidural

POST ANESTHESIA ASSESSMENT

Procedure completed as scheduled.
Anesthesia: General (ETT), Epidural
ASA Level: 3.0

Anesthesia Technique(s):
GENERAL (PRINCIPAL)
SPINAL

The Question

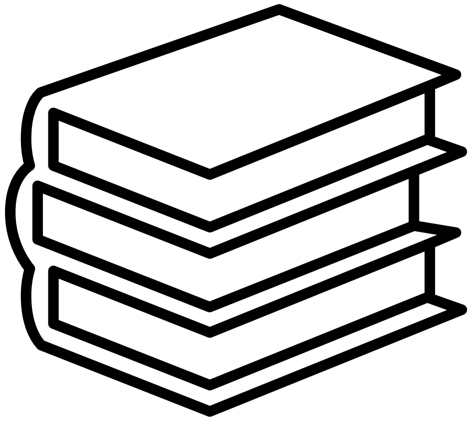
Can we leverage **natural language processing (NLP)** to improve our identification of **regional anesthesia (RA)** in VA surgical procedures?

Our Process

Developing the Framework

- Continued our **manual chart review** to ensure feasibility and determine the data structure
  - Where is the information stored?
  - Is it consistently recorded across sites and years?

- **Literature review**
  - What has been done prior?
  - Can we build on prior research?



The Framework

- Data Collection
- Text Extraction
- Text Cleaning
- Named Entity Recognition
- Text Classification
- Information Extraction
- Evaluation and Validation

Data Pre-Processing

Algorithm Development

Step 1: Data Collection

- First, identify and collect the text data.
  - We identified our text data in the **TIU domain.**
- We were very **broad** at first.
  - This is an iterative process, so this can be revised later.
- We started by tabling **TIUStandardTitle** and **TIUDocumentDefinition** values and pulled specific notes that matched an operative report or anesthesiology note.
- We then looked at the notes by Sta3n and year of surgery to ensure consistency.

Data Collection

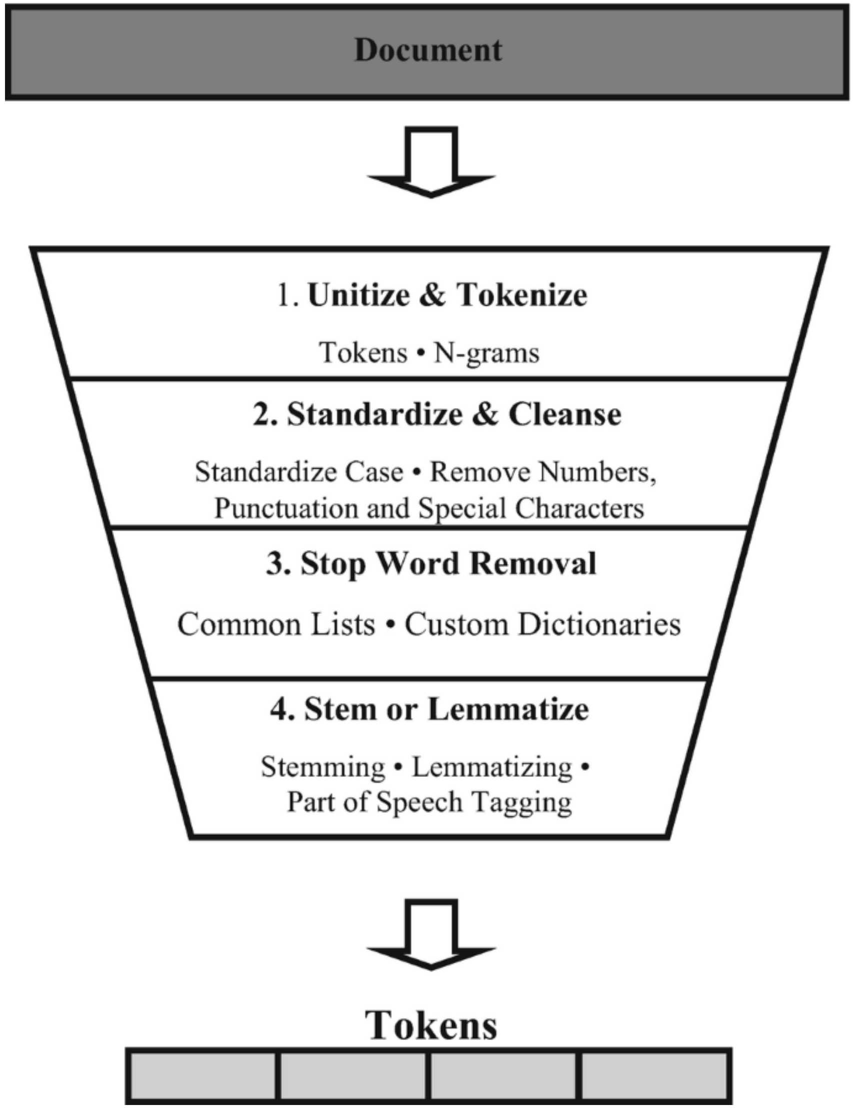
- **All sites** had an **Operation Report** and **Nurse Intraoperative Report** for all patients (TIUDocumentDefinition = 'OPERATION REPORT' OR 'NURSE INTRAOPERATIVE REPORT')

- The type and contents of **anesthesiology notes** (TIUStandardTitle = 'ANESTHESIOLOGY NOTE' OR 'ANESTHESIOLOGY PROCEDURE NOTE') varied by study site.
  - Some sites had detailed notes for RA procedures
  - Others used post-procedure notes to document anesthesia type
  - Still, some just uploaded PDF documents from their anesthesia record keeper.

Steps 2 & 3: Extraction and Cleaning

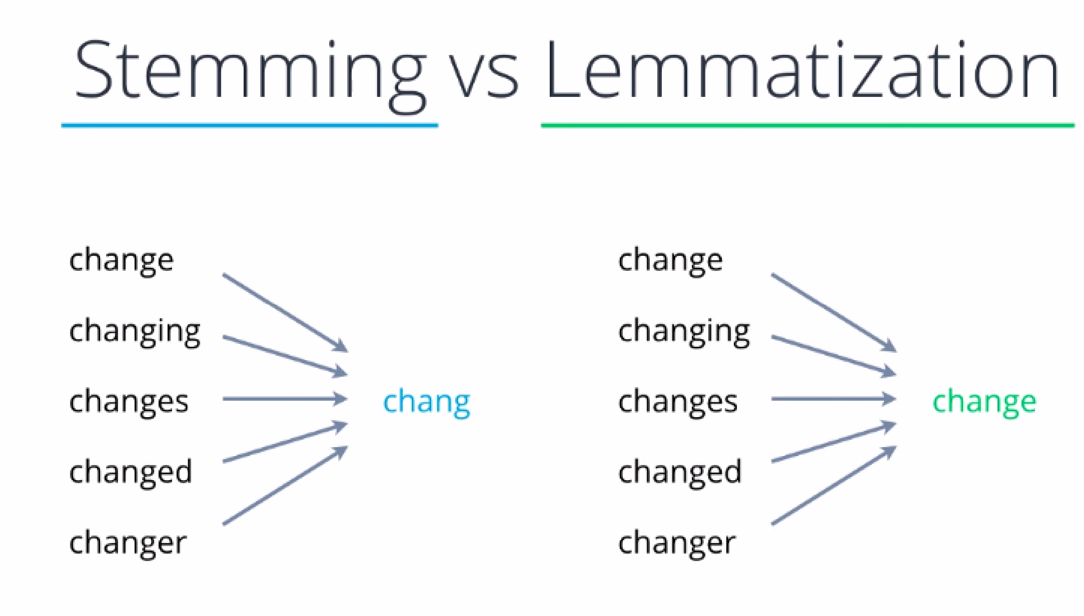
- The goal of pre-processing data is **text normalization and improving information retrieval tasks**.

- Text Cleaning and Standardization
  – Change all text to lowercase
- We also **identified starting points** in the text for structured text fields like
  – "ANESTHESIA:"
  – "Anesthesia Technique(s):"



Other Pre-Processing Options

- **Stemming**
  - Obtain the base or root form
- **Lemmatization**
  - Remove only inflectional endings.
  - Often more accurate than stemming and produces actual dictionary words, it can also be more resource-intensive



Steps 4 & 5: Named Entity Recognition and Classification

- Named Entity Recognition is an NLP method that identifies and classifies important information in text, also known as named entities.

- Our code was developed to **identify keywords and dates**
  - We looked for basic keywords like “epidural”, “spinal”, “regional anesthesia”, or “nerve block”
  - As we evaluated our code, we extended the searches to more specific variations , like “interscalene nerve block” or common misspellings like “interscaline nerve block”.

Step 6: Information Extraction

POST ANESTHESIA ASSESSMENT

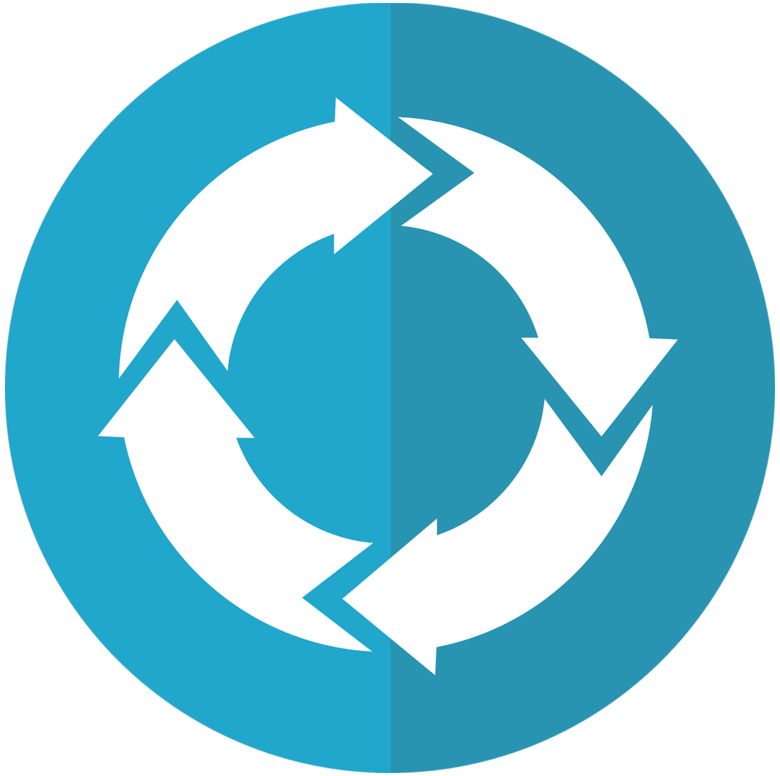
Procedure completed as scheduled.
Anesthesia: General (ETT), Epidural
ASA Level: 3.0

Regional = 0
Epidural = 1
Spinal = 0
Block = 0

- References to RA were identified in the data using **binary codes (0/1)**
- The initial dataset contained **one row per note,** so there were multiple rows per procedure
- After we had settled on an algorithm, we **summarized our information to the procedure level**.

Step 7: Evaluate and Repeat

- A random **sample of 250 cases** was manually reviewed to refine the algorithm's accuracy.
- If the overall accuracy of the algorithm in the sample was <95%, then we repeated the process.
- In the end, **96.4%** of RA identified by the algorithm were confirmed with manual chart review.



Our Results

Regional Anesthesia & Pain Medicine

Brief technical report

Use of natural language processing method to identify regional anesthesia from clinical notes

Laura A Graham [1], Samantha S Illarmo [2], Sherry M Wren [3, 4], Michelle C Odden [2, 5] and Seshadri C Mudumbai [6, 7]

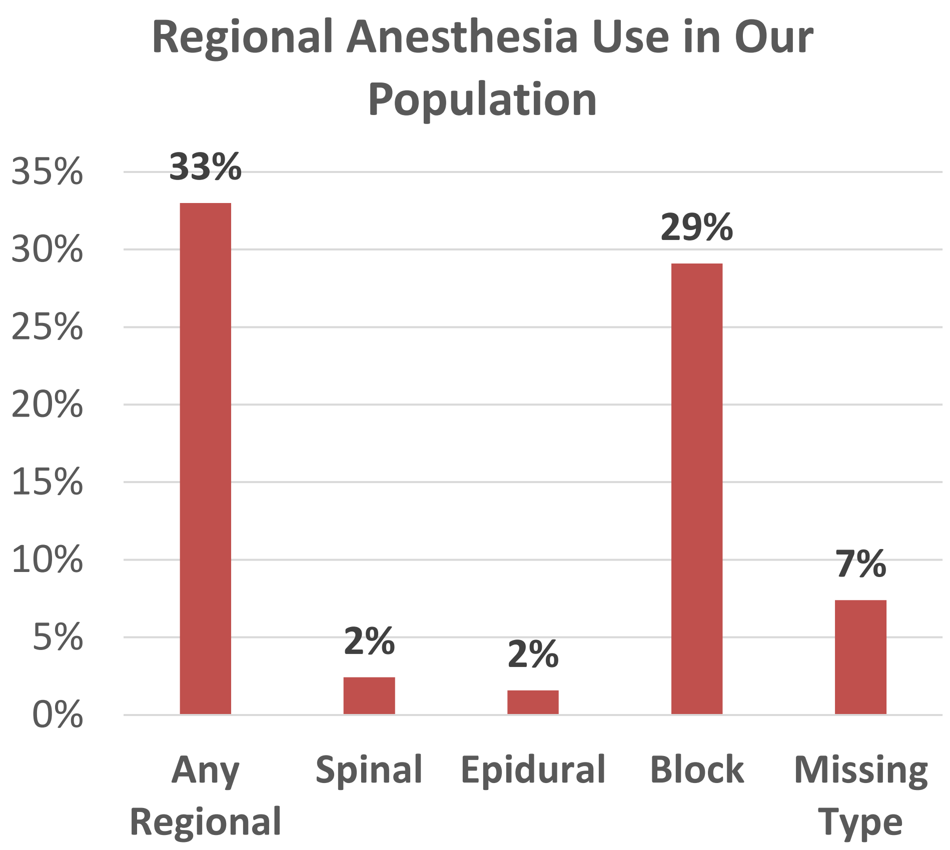
Correspondence to Dr Laura A Graham, Health Economics Resource Center, VA Palo Alto Health Care System, Menlo Park, California, USA; lagraham@stanford.edu

Abstract

Introduction Accurate data capture is integral for research and quality improvement efforts.

Results

- **27,713** surgical procedures
- **33.0% (n=9,154)** involved at least one RA procedure
- 88.1% of RA procedures involved a block
- RA was more common among younger patients ($p<0.01$) undergoing longer inpatient procedures, with significant variation by surgical specialty (0.5%–69.9%).



Algorithm Validation

- To assess the accuracy of our algorithm to traditional methods, we compared the algorithm's results to **the documentation of RA in the CDW** (gold standard).

- We used **sensitivity**, **positive predictive value**, and **accuracy** to describe and quantify agreement between the two sources of information.

- The **false negative rate** was used to assess the completeness of information in the referent.

Validation Against CDW

- **96.6%** of CDW RA+ cases were identified with the NLP algorithm.
- Only **48.6%** of NLP RA+ cases were also CDW RA+.

		CDW (Referent)	
		+	-
NLP	+	4,450	4,704
	-	156	18,403

Sensitivity	**0.966**
Positive predictive value	**0.486**
Accuracy	**0.825**
False negative rate	**0.008**

Overlap with CDW

Sensitivity Analysis

- To aid in generalizability outside of the VA medical record, we also developed a logistic regression model for text classification using the clinical note data.
- We used the **“bag-of-words”** approach, incorporating all text in the clinical notes, and applied a term frequency-inverse document frequency matrix
  - Each row represents a document,
  - Each column represents a word or phrase containing up to three words.
- The model was **trained on a 75% sample**, and the remaining 25% was used to test the model’s performance.

Sensitivity Analysis

Measures of agreement			
	Sensitivity	**Positive predictive value**	**Accuracy**
Trained on the CDW data	0.975	0.867	0.855
Trained on the NLP-based algorithm	0.951	0.848	0.860

Some Caveats

- Missing anesthesia type
  - 85% of CDW regional anesthesia cases were missing the type of regional anesthesia.
  - 32% of these reports contained only scanned PDF documents from the facility's perioperative management system, which we were unable to use.
  - The amount of missing information raises concerns about the completeness and accuracy of data that can potentially be used for research and quality improvement both in the future and historically.
  - Not only can missing data bias research results if these data are not missing at random, but missing data leads to underestimates of prevalence.

Summary

- **33.0% (n=9,154)** of patients in our sample received RA.
- However, **50% of these cases were not documented** in one of the most commonly used administrative VA data sources, the CDW Surgery Domain.
- 85.1% of CDW RA+ cases were **missing the type of RA.**
- Further, **nerve block documentation** in VA data is minimal, limited to only one information source, highlighting an important area for improvement.

Questions?

Check out the Podcast!

RAPM Focus

BMJ Journals

Subscribe Log In Basket search

Regional Anesthesia & Pain Medicine

Latest content Current issue Archive

Home / Online First

Article Text

Article info

Citation Tools

Share

Brief technical report

PDF

PDF + Supplementary Material

Use of natural language processing method to identify regional anesthesia from clinical notes

Laura A Graham [1], Samantha S Illarmo [2], Sherry M Wren [3, 4], Michelle C Odden [2, 5] and Seshadri C Mudumbai [6, 7]

Correspondence to Dr Laura A Graham, Health Economics Resource Center, VA Palo Alto Health Care System, Menlo Park, California, USA; lagraham@stanford.edu

Abstract

Introduction Accurate data capture is integral for research and quality improvement efforts.

Link to the article