Detection of prone positioning in hospitalized COVID patients using NLP

Patrick R. Alba, Jose Posada, Jason Weatherald, Kristine E. Lynch, Annie Bowles, Nigam Shah, Olga V. Patterson, Scott L. DuVall*, Evan Minty*

Background

The majority of patients diagnosed with Coronavirus Disease–19 (COVID-19) are asymptomatic or have mild disease, yet approximately 14% develop more severe disease including hypoxemic respiratory failure and/or Acute Respiratory Distress Syndrome (ARDS) (1).

Prone positioning is a life-saving intervention for mechanically ventilated patients with moderate-severe ARDS with a number needed to treat of 6 to save one life (2–4). Although there is strong evidence for prone positioning in mechanically ventilated patients with ARDS, there is a lack of robust evidence for its use in awake non-intubated patients.

Prone positioning patients who are awake and not intubated is a novel but experimental idea. It is a non-pharmacological therapy, which would be low-cost and simple to apply, even in low resource settings. Despite the absence of evidence for clinical efficacy, awake prone positioning has been widely adopted in many centers during the COVID pandemic, and was conditionally endorsed by organizations such as the World Health Organization and the Intensive Care Society (5) (6).

There are no large-scale, real-world data on the prevalence of awake prone positioning among hospitalized patients with Covid-19. In this proof-of-concept study, we aimed to use natural language processing (NLP) to investigate the incidence of awake prone positioning in a large cohort of hospitalized patients with Covid-19 across multiple institutions.

Methods

Study Subjects

The study was conducted within the US Department of Veterans Affairs (VA) Informatics and Computing Infrastructure (VINCI), where NLP methods for prone detection were developed in collaboration with the study authors. As of time of writing, proof of principal deployment of the NLP pipeline developed in VINCI has been established within the Stanford Health system data warehouse (STARR), against n = 10 clinical notes. We leveraged existing cohort definitions for COVID positive hospitalized patients developed in the OMOP data model for the OHDSI CHARYBDIS study (7). For the purposes of this proof of principle study, the CHARYBDIS cohort logic informed a query against a non OMOP VA VINCI database. At Stanford, the CHARYBDIS definition was used in the STARR OMOP repository. The current state and future work relevant to completing a proning incidence study using VINCI and STARR are summarized in table 1.
<table>
<thead>
<tr>
<th>Cohort</th>
<th>Current State and future work in VINCI</th>
<th>Current state and future work in STARR</th>
</tr>
</thead>
</table>
| COVID positive Hospitalized    | • CHARYBDIS logic, implemented in non OMOP VINCI DB in this submission (C\textsubscript{VINCI})  
• Future work: implement in VINCI-OMOP                                                                                                                                          | CHARYBDIS cohort, implemented in OMOP (C\textsubscript{STARR})                                    |
| Treated with Prone Positioning | • Algorithm developed in VINCI, deployed against entirety of C\textsubscript{VINCI}  
• Performance evaluated in an adjudicated sample of 100 patients                                                                                                           | • Algorithm developed in VINCI, deployed as proof of principle against n=10 notes in C\textsubscript{STARR}  
• Future work: deploy against entirety of C\textsubscript{STARR} for external performance estimate of original algorithm  
• Future work: use feedback from STARR performance to iterate algorithm |

Table 1: Current state of cohort development and implementation in VINCI and STARR relevant to the implementation of a proning incidence study

**NLP for detection of patients exposed to prone positioning and ‘intent to prone’**

Clinical notes for COVID patients were often found to contain recommendations regarding prone positioning in awake patients (e.g. notes indicating that proning was encouraged, or that education was provided). This reflects an ‘intention to prone’, regardless of the patient’s subsequent exposure to prone positioning. Subsequent tolerance of awake prone positioning appears to be variable in real world data, similar to prospective trials of awake proning in COVID patients(8).

Granularity at the level of ‘intention to prone’ is therefore of interest to future efforts to characterize exposed, intent to expose, and non-exposed individuals. To permit this flexibility in subsequent analysis, we decided to roll up all mentions of proning into the following categories at a term incidence level:

- **Treated**: positive mention indicating the patient was exposed to prone positioning.
- **Intent to treat**: proning appears to have been encouraged, with no clear negation or indication as to its tolerability.
- **Not treated**: typically reflects an inability to prone, or intention to prone that is negated.

At a patient / admission level, a patient may have several mentions documented across time that reflects any combination of the above categories. Alternatively, they may have no documentation identified related to proning attempts or positioning, or there may be ambiguity in the documentation that doesn’t allow us to clearly classify some of the mentions. At an admission level, we rolled up mentions during admission to the following categories:

- **Treated**: any positive mention affirming prone positioning, regardless of other mentions (e.g. of intent or negation).
- **Intent to treat**: intent only (no mentions affirming prone position, or refusal)
- **Not prone**: explicit refusal or inability to prone.

A set of theoretical patient timelines illustrating how text instances would be classified at a patient admission level is shown in figure 1.
NLP Pipeline
We first identify all instances of a prone term in a patient's notes (e.g. “Prone”, “proning”, “proned”, etc). Next, the most frequent and irrelevant terms are identified and excluded (e.g. ‘Accident-prone’, ‘prone to falls’ etc.). All remaining terms are then classified to the categories above at an instance-level using concept matching dictionaries and rule-based patterns. The NLP system was built using the Apache Unstructured Information Management Architecture Asynchronous Scaleout (UIMA-AS) (9) and the libraries and tools contained in the VA Leo framework (10). At time of writing, we have successfully deployed the pipeline to annotate a small set of notes of COVID positive admitted patients represented within Stanford STARR as proof of feasibility.

![Diagram](image)

**Figure 1**: Theoretical patient timeline examples demonstrating classification at a patient / admission level. Note that oxygenation requirement windows are shown for illustrative purposes, (mechanical ventilation detection is implemented at the VA site).

Results
Admission level incidence of proning exposure (treated, intent to treat, not treated) are shown in table 2. We also report on ‘Treated’ patients both within and outside of windows in time when they were being mechanically ventilated as determined by a separate VA NLP classifier and data from the VA COVID-19 Shared Data Resource. As of time of writing, admission level performance characteristics of the ‘Treated’ category have been assessed at the VA on a sample of 100 patients. Interobserver agreement at a term instance level revealed a Kappa of 0.73. Algorithm performance showed a sensitivity, specificity and positive predictive value of 84.09, 94.64, and 91.28 respectively. Algorithm performance is currently being assessed within Stanford STARR.
### Table 2: Admission level incidence proportion showing treatment, intent to treat, or ‘no definitive exposure’ to prone positioning as defined through NLP. Counts shown are reflective of data from March 2020 to May 2021.

<table>
<thead>
<tr>
<th>Site</th>
<th>Cohort</th>
<th>Admission level Proning Incidence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>COVID Hospitalized, mechanically ventilated at time of term incidence</td>
</tr>
<tr>
<td>VA</td>
<td></td>
<td>1,403</td>
</tr>
<tr>
<td></td>
<td>Treated: positive mention affirming prone positioning</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intent to Treat: Intent only Intent and not treated</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Not treated</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Prone term identified, no classification made</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No documentation identified</td>
<td></td>
</tr>
<tr>
<td>Stanford</td>
<td>(All admission level categories)</td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Conclusion**

Prone positioning of awake patients in respiratory distress has emerged as a potential therapeutic adjunct in COVID-19 respiratory disease. In real world data, there are unlikely to be structured data elements that capture the prone positioning of a patient. Our demonstration and use of NLP methods provides the first large-scale estimate of proning incidence, including awake proning, in the treatment of COVID patients. Future work will focus on completion of multi-institutional proning incidence study through large scale deployment within the Stanford STARR cohort. This will allow external validation of the original algorithm, as well as iteration that may improve generalizability. This will be followed by further work to better resolve ‘intention to treat’ and ‘not treated’ classes (many examples in the latter category reflect intention to prone and negation / refusal, which are of interest to reclassify as ‘intent to treat’). With a determination of the performance characteristics of those classes, we would hope to be able to further characterize patient populations exposed and not exposed to proning, which may help contextualize forthcoming clinical trial results in this area.
References/Citations


