Abstract:

Background: Data from Electronic Medical Records (EMR) is increasingly used to conduct retrospective chart reviews (RCHR) for clinical research purposes. Hospital IT departments are tasked with creating an algorithm to abstract desired data from EMRs.

Objectives: First, to demonstrate that initial data output can contain inaccuracies. Second, to devise practices that address these inaccuracies before the data acquisition process begins.

Methods: Data on extubation status and airway management devices had been collected between 2012 and 2014, using a data abstraction algorithm for the following patient populations: ASA Status 1&2 patients undergoing circumcision and ASA Status 3&4 patients undergoing tonsillectomy with or without adenoidectomy (T&A). A manual chart review was initiated when the data set was found to be inconsistent with clinical experiences. The frequency data obtained from each method was then compared.

Results: For extubation status in the study population, automated data abstraction produced 246 “unknowns.” 63.4% of these were identified through manual chart review. If output marked “none” for airway device in 115 cases. Manual chart review revealed that an airway device was used in 97.4% of these cases.

Conclusion: Automated data abstraction can produce incomplete and inaccurate data which can skew study results, lead to erroneous conclusions, and misguide future prospective studies. It is important for a trained clinician to audit the results. Close collaboration between the research and IT teams regarding specific fields to query, as well as the related data outputs, can help mitigate these inaccuracies. Our research team now creates a “blueprint” for subsequent data pulls to address these issues.

Methods:

- Data on extubation status and airway management device between June 2012 and 2014 was collected using automated data collection methods for two patient populations:
  - ASA 1&2 patients undergoing circumcision (n=1,346)
  - ASA 3&4 patients undergoing T&A (n=308)

- Exclusion criteria: patients over the age of 21

- A manual chart review was initiated when collected data appeared inconsistent with clinical experiences

- Frequency data obtained from each method was compared

- Sources of error were identified

- A blueprint was devised based on these sources of error to help increase the accuracy of automated data collection

Results:

- Initial data pull results yielded 115 cases with no airway device (“none”)

- Manual chart review revealed the data abstraction algorithm was outputting “none” when the queried field was blank, NOT when there was no airway device used

- 52 of these patients had an ETT, 53 had a LMA, 1 had monitored anesthesia care, and 6 had a mask, while only 3 cases remained unknown

- Initial data pull results indicated the extubation status for 246 cases was “unknown”

- Manual chart review revealed 356 of these to be awake (38), deep (6), not extubated (7), not intubated (104) and reintubated (1), (98) remained unknown

Conclusion:

- Automated data abstraction can produce incomplete and inaccurate data

- Sources of error were found to be:
  - The fields being queried
  - Related algorithm outputs
  - Terminology interpretation

- Close collaboration between research and IT teams regarding specific fields to query and the corresponding data outputs can help prevent erroneous data collection

- Create a blueprint to share with the IT team for upcoming data pulls to improve accuracy and efficiency

- Have data audited by a clinician

- Future analyses will include quantifying the efficacy and efficiency of the blueprint method on an upcoming data pull

References:


Figure 1 & 2: Comparing airway device and extubation status frequency percentages obtained from each data collection method

Figure 3: Blueprint exemplifying how to utilize the method to pull airway device and extubation status