Machine-Learning Approaches to Identifying Error

Eric Ford, PhD, FAAPM
Professor
University of Washington
Seattle, WA

AAPM 2018, Nashville, TN. Thur 10:00-12:00 am. 35 min talk + 5 min for questions. TH-CD-KDBRA1-1: Errors and Data Mining in EMR
Disclosures

• R18 HS022204-01
• NCI UG3 CA211310-01
Learning Objectives

• Understand existing and emerging methods for analyzing electronic medical record (EMR) data to quantify gaps in the safety and quality of care.
• Understand how the EMR can be used to continuously monitor the status of patients for clinically important endpoints.
• Appreciate the role of the EMR and human-computer interaction in contributing to error.
Case Study
SBRT Planning Error: MU doubled

- 8 Gy x 5 liver plan
- Plan printed
- Planner is in a rush. Clicks “No plan normalization”
- Printed plan looks OK (% isodose lines)
- Plan QA. Not caught.
- Caught by RTT just prior to beam on
Existing Tools for Finding Errors and Hazards
FMEA

Failure Mode and Effects Analysis

Incident Learning
Assuming the audience is somewhat familiar with FMEA and ILS, but if not here are some references.
This is from a study in our group on SBRT which looked at all the events that can happen versus what was observed through incident learning.

Yang et al. Med Phys, 42(6), 2777-2785, 2015
Incident learning and Risk analysis are complementary

All events 56

23 events (41%) NOT observed with incident learning

Incident Learning 33 events

Yang et al. Med Phys, 42(6), 2777-2785, 2015

Link to FMEA and resident education
The Incident Reporting System Does Not Detect Adverse Drug Events:
A Problem for Quality Improvement

Audit: 54 Adverse Drug Events
3 incident reports (6%)
A different approach:
Mining the Health Record
THE NATURE OF ADVERSE EVENTS IN HOSPITALIZED PATIENTS

Results of the Harvard Medical Practice Study II

Lucian L. Leape, M.D., Troyen A. Brennan, M.D., J.D., M.P.H., Nan Laird, Ph.D.,
Ann G. Lawthers, Sc.D., A. Russell Localio, J.D., M.P.H., Benjamin A. Barnes, M.D.,
Liesl Hébert, Sc.D., Joseph P. Newhouse, Ph.D., Paul C. Weiler, LL.M., and Howard Hiatt, M.D.

New England Journal of Medicine, 324(6), 377-384, 1991

Very expensive
Harvard Health Study

1984, 51 hospitals in NY
n=30,195 charts randomly sampled (in-patients)

Method:
- Screen for adverse events
- “unintended injury cause by medical management that resulted in measureable disability”
- RN screen (18 criteria)
- 2 MDs review & evaluate, rate disability

1,133 adverse events (3.7%)
Errors in management in 58% of these

Very expensive
This lead to the landmark IOM study. But note: the process used in the Harvard Health study relies on records and extremely labor intensive and expensive.

Institute of Medicine (IOM) 1999

44,000 to 98,000 people die in US hospitals each year due to preventable errors.
Automatic Error Mining

“Trigger Indicators”
Key thing: the automatic mining relies on an HER. Our EMR(s) have a wealth of information, waiting to be extracted and used.
Growth of the EHR ... 97% of hospitals use an EHR now. Grew fast. How do you “mine” the EMR? One method is to randomly sample charts ...
Audit: 54 Adverse Drug Events
3 incident reports (6%)

565 citations. Example automated signals: ordering known antidotes (e.g. nalaxone hydrochloride Rx, vitamin K), lab tests for serum drug levels
Error indicators in the OIS

A field is “hidden”

A prescription is changed

Document is voided

{}
Training and Validation

• 15 potential indicators (MQ)
• Correlate: safety event per patient
• LR model
train n=2210, test n=949

AUC 0.653
Accuracy 90.7%

Hartvigson et al. pending 2018
Can we predict for or identify error before it happens?
“Correlates” Studies

• Treatment type “x” is more error prone
# “Correlates” Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Higher error rates for ...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morganti et al. 2008</td>
<td>More complex treatments</td>
</tr>
<tr>
<td>Dominello et al. 2015</td>
<td>Weekend cases and “sim and treat”</td>
</tr>
<tr>
<td>Walker et al 2015</td>
<td>Time, # fractions, more Rx items, beam duration</td>
</tr>
<tr>
<td>Elnahal et al. 2016</td>
<td>Peds, H&amp;N, breast, pts on protocol, IMRT/IGRT</td>
</tr>
<tr>
<td>Gensheimer et al. 2016</td>
<td>Peds</td>
</tr>
<tr>
<td>Judy et al. 2017</td>
<td>H&amp;N, IMRT, daily IGRT</td>
</tr>
</tbody>
</table>

References in notes on handouts

Conclusion ... studies vary, institutional data, no one obvious predictor.
Probabilistic Network for Error Detection

Bayesian Network

Probabilistic Network

- Brain
- 300 cGy/#
- 5400 cGy
- 2D
- Definitive

X
Validation Through Simulated Error

Sensitivity

1-Specificity

Brain
Breast
Lung

Bayes Net

Human observer

Sensitivity

1-Specificity

Expanded network

Kalet et al. in preparation 2018
SBRT Planning Error: MU doubled

- 8 Gy x 5 liver plan
- Plan printed
- Planner is in a rush. Clicks “No plan normalization”
- Printed plan looks OK (% isodose lines)
- Plan QA. Not caught.
- Caught by RTT just prior to beam on

Of note here – this error might NOT have been reported to an ILS, but either the trigger indicator or the probabilistic network could pick it up.
These tools could be used to probe your ILS and safety culture, e.g. identify problems and see if they are reported into the ILS. That gives some indicator of the ‘health’ of the clinical safety program, a “grade”.

Conclusions

- EMR: automated identification of errors
- *Before* they reach the patient
- Complement to ILS and other tools
Acknowledgments

Lulu Jordan, (BS)RTT
Lora Holland, (BS)RTT
Patty Sponseller, CMD
Matt Spraker, MD
Alan Kalet, PhD
Mark Phillips, PhD
Matt Nyfot, PhD
Jing Zeng, MD
Ralph Ermoian, MD
Gabrielle Kane, MD

Medical Physics Residents!
Thank You!

eford@uw.edu