The University of Utah
College of Nursing

Representation of Smoking-related Concepts in an Electronic Health Record

Mollie R. Poynton, PhD, APRN[1,2],
Lewis Frey, PhD[2], Hongying Tang[2]

[1]University of Utah College of Nursing, Salt Lake City, UT
[2]University of Utah School of Medicine, Department of Biomedical Informatics, Salt Lake City, UT

The University of Utah
Biomedical Informatics

Purpose/ Aims

(1) Assess the feasibility of a knowledge discovery project related to smoking cessation using warehoused data from a group of outpatient clinics, AND

(2) Map strengths and gaps in the representation of smoking cessation related concepts in a common electronic health record system (EHR), EpicCare EMR.

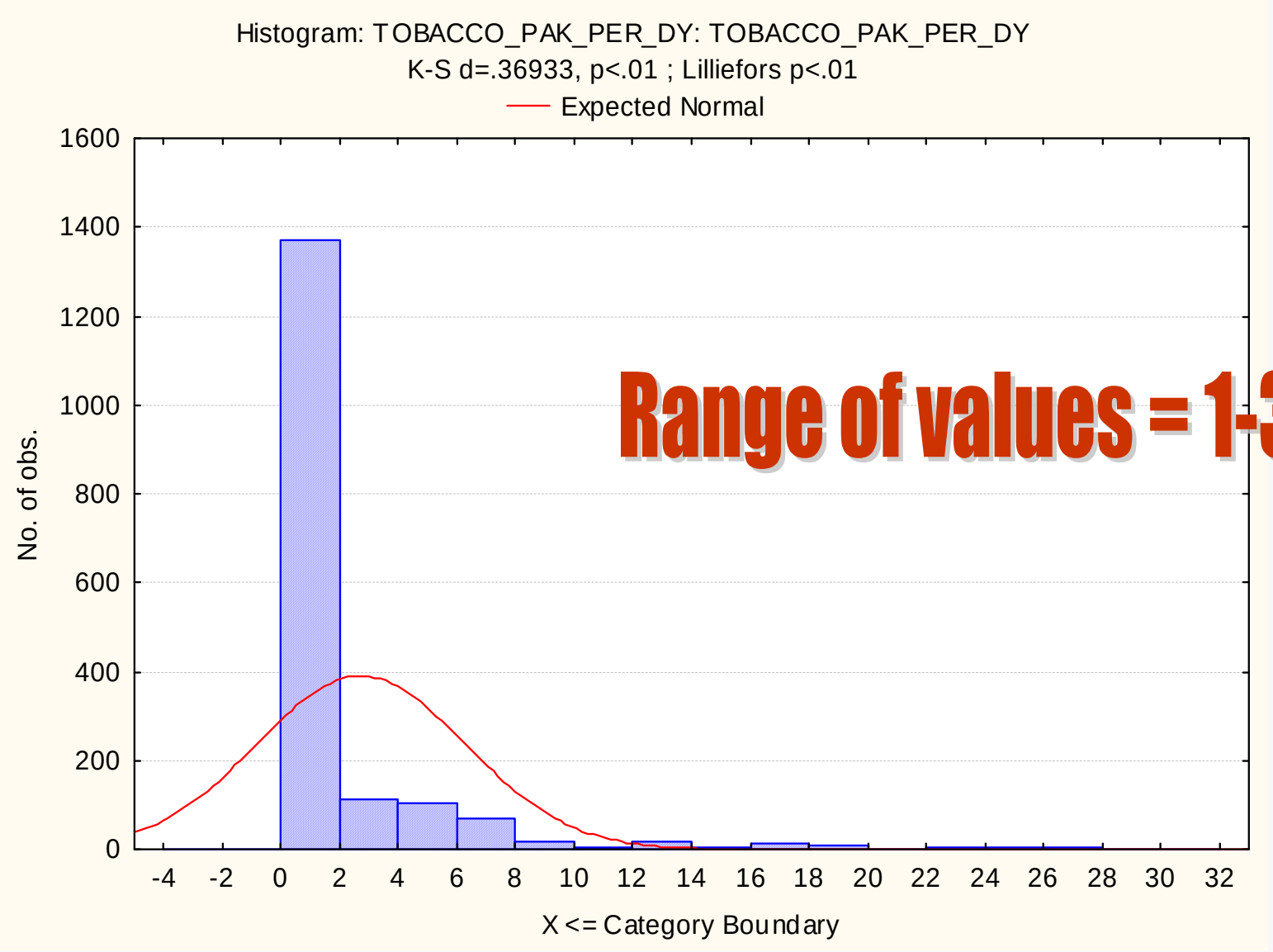
Data

- Source system: EpicCare EMR[4]
- Group of community outpatient clinics in the U.S. intermountain west.
- Data since 2000 (over six years of data).
- Random sample of 3000 electronic records:
  - Patients with a completed social history record
  - Either current or former smoker

Intensity and Duration of Smoking

Tobacco Status	***Current User (n=1693)***	***Former User (n=1307)***
Tobacco Packs Per Day	n=1220	n=521 **39.9%**
Mean (S.D.)	2.54 (3.48)	3.15 (3.69)
Tobacco Years	n=1012	n=542 **41.5%**
Mean (S.D.)	15.71 (13.28)	15.18 (12.86)

Packs Per Day: Distribution



Word frequencies: Tobacco Comment

	Frequency	Number of documents	Stemmed word
age	185	171	age
ago	[illegible]	[illegible]	ago
day	199	196	day
quit	561	[illegible]	quit
since	124	123	sinc
smoke	204	196	smoke
year	157	145	year
yr	178	167	yr

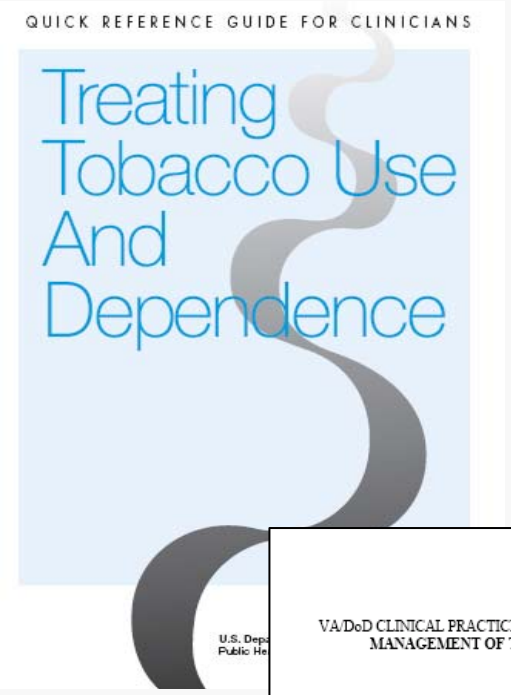
Words used to describe length of time...

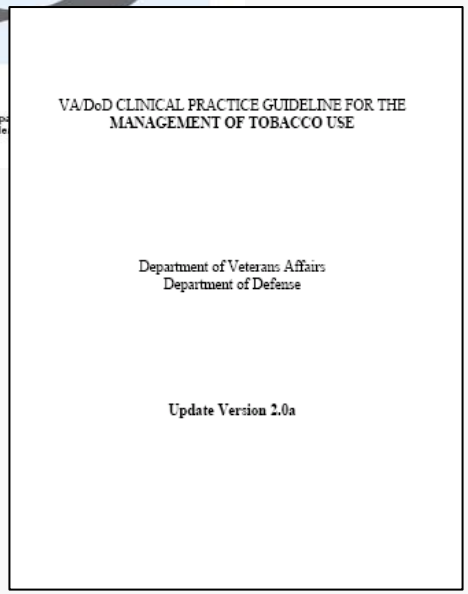
Concept Mapping

U.S.P.H.S. Treating Tobacco Use and Dependence[5]

Va/DOD Clinical Practice Guideline for the Management of Tobacco Use[6]

Additional concepts identified in review of the literature[1]





Concept Representation: Strengths

- Prescribed medications
- ICD9 code “tobacco use disorder”[8]
- Willingness to quit
- Quit date
- Presence and type of tobacco use
- Arrangement of follow up
- Demographic information
- Family history/ pedigree

Concept Representation: Gaps

- Advice/counseling to quit
- Educational intervention
- Self-help materials
- Referral to external counseling/ treatment programs
- Type/ characteristics of program
- OTC nicotine replacement
- Social support

Conclusions

- Coded attributes capture some (but not all) concepts relevant to smoking cessation.
- Data assay indicates substantial data quality issues in coded data elements related to smoking cessation.
- Free text patient notes necessary to assess/ describe primary care management of smoking/ tobacco use. (Coded data *not* sufficient).

Discussion

- Data quality related to electronic health record use
- Recent software upgrade
- Extraction of smoking cessation related concepts from free text feasible (Hazlehurst, Sittig, Stevens, et al)[9]
- Need to move patient notes into EDW (enterprise data warehouse)!

References/ Acknowledgements/ Contact Information

1. Fayyad UM, Piatetsky-Shapiro G, Smyth P. From data mining to knowledge discovery. In: Fayyad UM, Piatetsky-Shapiro G, Smyth P, Uthurasamy R, eds. *Advances in Data Mining and Knowledge Discovery*. Menlo Park, CA: AAAI Press/ The MIT Press; 1996:1-34.
2. Poynton MR, McDaniel AM. Classification of smoking cessation status with a backpropagation neural network. *J Biomed Inform.* Mar 24 2006.
3. Poynton M. *Classification of smoking cessation status with machine learning methods*. Bloomington, IN: Graduate School, Indiana University; 2005.
4. *EpicCare EMR* [computer program]. Version; 2000-2006.
5. Epic Systems Inc. Clarity Enterprise Reporting Spring 2005 Data Dictionary.
6. Fiore MCB, W.C.; Cohen, S.J., et al. . *Treating Tobacco Use and Dependence.Clinical Practice Guideline.* Rockville, MD: U.S. Department of Health andHuman Services, Public Health Service; 2000.
7. Veterans Administration DoD. *VA/DoD clinical practice guideline for the management of tobacco use.* Washington (DC) Department of Veteran Affairs; 2004 June 2004.
8. International Classification of Diseases, Ninth Edition, Clinical Modification (ICD-9-CM).
9. Hazlehurst B, Sittig DF, Stevens VJ, et al. Natural language processing in the electronic medical record: assessing clinician adherence to tobacco treatment guidelines. *Am J Prev Med.* Dec 2005;29(5):434-439.

Thanks to Cheri Hunter and Ken Gondor from the University Health Care Data Resource Center in Salt Lake City, Utah for their advice and assistance.

Funded by the University of Utah College of Nursing Research Committee.

Conducted in cooperation with the Utah Health Research Network (UHRN).

Mollie R. Poynton PhD, APRN,

Assistant Professor, University of Utah College of Nursing

10 S 2000 East, Salt Lake City, UT 84112-5880, USA

Phone: 801.585.9740, E-mail: mollie.poynton@nurs.utah.edu