

POSTER ABSTRACTS

11th Annual HMO Research Network Conference

April 4-6, 2005 Santa Fe, NM

Information Technology 43

Identifying Depression Cases From Clinician Notes Using Natural Language Processing

LR Fischer, WA Rush, JC Kluznik; AL Crain, AM Hanson, JE Duncan, PJ O'Connor

Background: In research studies in health care systems, depression cases can be identified through diagnosis (ICD-9) codes. However, because most encounters have only one or two ICD-9 codes, it is likely that depression often is under-coded. Clinician notes offer an alternative source for identifying depression. While traditional chart audits are time-consuming and costly, natural language processing (NLP) tools can be used to enhance case identification where clinician notes are available in electronic form.

Methods: An ACCESS-based computer tool was developed to elicit key terms in open-ended text data. The tool has the capability to track key terms (words, parts of words or phrases), link these terms with identifiable clinician notes, and create a list of phrases with the key term imbedded. A qualitative analysis was used to identify key terms and exclusion terms. To test the viability of this approach, a random sample of 1,000 adult HMO patients with no ICD-9 code for depression was selected in 2001, along with 100 patients with a depression ICD-9 code. Using a list of phrases that included the imbedded term "depression," a researcher coded all cases with an appropriate indication of depression, excluding inappropriate uses of the term ("denies depression", "depression fracture", "no history of depression"...).

Results: Of 1,000 patients with no depression code, 133 (13%) had no encounter. Among the 867 remaining cases, depression was noted for 51 patients—that is, 5% of this sample had at least one depression-related encounter but would have been missed by ICD-9 codes alone. The term "depression" was used in 91 of 100 cases with an ICD-9 code for depression.

Conclusions: ICD-9 codes for depression miss substantial numbers of cases. The NLP tool provides an efficient way to enhance the identification of cases. This tool could also be used for other diagnoses. Further, as more and more medical providers establish paperless offices, NLP will be increasingly useful in both research studies and clinical practice.

This study was funded by a grant from NIH/NIDDK: #DK066050