Derivation of Malignancy Status from ICD-9 Codes

[1]Mark G. Weiner, M.D., [1]Alice Livshits, [2]Carol Carozzoni, Pharm.D., [2]Erin McMenamin, [3]Gene Gibson, Pharm.D., [2]Alison W. Loren, M.D., [2]Sean Hennessy, Pharm.D., M.S.C.E.
[1]Division of General Internal Medicine, [2]Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania School of Medicine, Philadelphia, PA 19104
[3]Department of Pharmacy Services, Hospital of the University of Pennsylvania

To assess the severity of illness of oncology patients, it is necessary to distinguish patients with a single primary tumor from patients with metastatic disease occurring at a secondary location remote from the primary site. We developed a ranked list of cancer groupings and an algorithm that could distinguish patients with primary and metastatic cancer even if no specific code for secondary cancer was recorded. In patients with metastatic disease, the algorithm should also distinguish the primary site from the secondary site.

INTRODUCTION

The ICD-9 dictionary includes a specific range of codes that represent primary and secondary cancers, however, empiric evidence shows secondary cancer codes are vastly underutilized. Frequently, the secondary cancer is incorrectly coded as if it were a primary tumor in the new location. We developed a ranked set of cancer categories that can be mapped to cancer ICD-9 codes recorded for each patient of a cohort. We hypothesized that patients with cancer codes that span more than one category would have metastatic disease in which the primary site would have the lower ranking and the secondary site(s) would have the higher ranking.

METHODS

We subcategorized the oncology section of the ICD-9 dictionary into anatomically related units as described in the table. The table was sorted such that cancers with lower ranking were more likely to be primary tumors, and cancers with higher ranking were more likely to be secondary tumors. In this manner, a patient with codes for lung cancer alone would be judged as having primary lung cancer, but a patient with codes for both colon and lung cancer would be considered to have colon cancer metastatic to the lung. We examined the ability of this algorithm to appropriately categorize patients into metastatic and non-metastatic cancer categories. We created a database of all inpatient and outpatient cancer codes for a set of 70 oncology patients without leukemia/lymphoma who were receiving ambulatory chemotherapy and were admitted within 30 days of chemotherapy. Chart review using a conservative measure that required explicit mention of metastatic disease was used as a gold standard. This definition favored the gold standard's assessment of non-metastatic disease

RESULTS

Of the 70 patients analyzed, 42 had metastatic disease according to the algorithm of which 32 had metastatic disease by chart review. Of the 28 patients without metastatic disease according to the algorithm, 24 also had no metastatic disease by chart review. The overall concordance between the methods was 80%.

CONCLUS ION

The table provides an anatomical grouping of cancer types that can be mapped to cancer diagnoses recorded for a patient. The proportion of patients having codes that span more than one group can be used to estimate the prevalence of metastatic cancer in the population.

Rank	Cancer Group	ICD-9 code ranges
1	Melanoma	172*
2	Breast	174*, 175*, 239.3
3	Colon	153*, 154*, 235.2
4	Gyn	180*, 182*, 183*, 184*, 236.1, 236.2
5	Prostate	185* 236.5
6	Testes/Male GU	186*, 187.3, 187.4, 187.9, 236.4, 236.6
7	Head and neck	140-149.9, 160*, 161*, 162*, 195.0
8	Urinary Tract	188*, 189*, 236.7, 236.91, 239.4, 239.5
9	Non-melanomatous skin cancer	173*, 238.2
10	Non-colon GI	150-152.9, 155-159.9, 235*, 239.0
11	Lung	162*, 235.9, 239.1
12	Brain	190-192.9, 237.5, 237.6, 239.6
13	Bones/soft tissue	170*, 171*, 238.1, 238.2
14	Endocrine	193, 194*, 237.0, 237.4, 239.7
15	Pleura/mediastinum	163*, 164*
16	Non-specific site	195*,199*,238.8,238.9 239.8, 239.9
17	Lymph node spread	196*
18	Secondary cancer	196*, 197*