C25

Multi-Feature Method for Ascertaining NET Cases in a Clinical Data Warehouse

Marie Synnestvedt PhD[1], Bonnie Bennett RN, BSN[1], Hillary Faust MD[1], Mark Weiner MD[1], David C. Metz MD[1]

[1]The Perelman School of Medicine at University of Pennsylvania, Philadelphia PA 19104

Background: A multidisciplinary program should be the standard of care for neuroendocrine tumors (NETs) but care is often fragmented across multiple specialties. We developed a method to ascertain how many NET patients existed in our health care system in order to establish a registry for future clinical care and research.

Methods: A knowledge discovery approach was undertaken to identify a trial set of criteria by examining terminologies within electronic NET resources such as NANETS and SNOMED and performing statistical and text analysis of a sample of 364 known NET cases (128 GI NETs and 236 pheochromocytoma/paragangliomas). The criteria selected included laboratory (any abnormal 5HIAA, chromogranin A, epinephrines, metanephrines, norepinephrines, pancreatic polypeptides, serotonin, or VMA), pathology (positive keywords in diagnosis), radiology (non-negative MIBG or Octreoscan), medications for pheochromocytoma, or ICD9 codes (for genetic diagnoses such as MEN). These criteria yielded 5262 potential NET cases going back to 1999. A 10% random sample of 302 cases was selected from 3017 unclassified cases with activity since 2008 for review and classification of true NET status.

Results: The search method retrieved 90% of known NET cases. Missed cases were primarily older cases that predated searchable electronic data or that were diagnosed using historical terminology such as glomus tumor. The true case rate in the random sample was 132/302 (44%) – 56 GI NETs, 14 phaeochromocytoma/paragangliomas, 27 lung carcinoids and 35 other NETs (170 were not NETs).

The least specific univariate criteria were the pathology keyword synaptophysin or abnormal metanephrines or norepinephrines. Removal of cases having single non-specific criteria increased the true positive rate to 127/226 (56.2%) with 3.7% missed true positives.

Conclusion: This methodology can be improved with further modeling, and in its current form can help institutions identify a larger cohort of previously unidentified patients with NETs who are suitable for inclusion in a registry for follow-up, management, and improved research capabilities.