Original Article

IDENTIFICATION OF POTENTIAL MARKERS FOR CUSHING DISEASE

Michael S. Broder, MD, MSHS[1]; Eunice Chang, PhD[1]; Dasha Cherepanov, PhD[1]; Maureen P. Neary, PhD, MS[2]; William H. Ludlam, MD, PhD[2]

ABSTRACT

Objective: Cushing disease (CD) causes a wide variety of nonspecific symptoms, which may result in delayed diagnosis. It may be possible to uncover unusual combinations of otherwise common symptoms using International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes. Our aim was to identify and evaluate dyads of clinical symptoms or conditions associated with CD.

Methods: We conducted a matched case-control study using a commercial healthcare insurance claims database designed to compare the relative risk (RR) of individual conditions and dyad combinations of conditions among patients with CD versus matched non-CD controls.

Results: With expert endocrinologist input, we isolated 10 key conditions (localized adiposity, hirsutism, facial plethora, polycystic ovary syndrome, abnormal weight gain, hypokalemia, deep venous thrombosis, muscle weakness, female balding, osteoporosis) with RRs varying from 5.3 for osteoporosis to 61.0 for hirsutism (and infinite RR for localized adiposity). The RRs of dyads of these conditions ranged from 4.1 for psychiatric disorders/serious infections to 128.0 for hirsutism/fatigue in patients with versus without CD. Construction of uncommon dyads resulted in further increases in RRs beyond single condition analyses; for example, osteoporosis alone had an RR of 5.3, which increased to 8.3 with serious infections and to 52.0 with obesity.

Conclusion: This study demonstrated that RR of any one of 10 key conditions selected by expert opinion was ≥5 times greater in CD compared to non-CD, and nearly all dyads had RR≥5. An uncommon dyad of osteoporosis and obesity had an RR of 52.0. If clinicians consider the diagnosis of CD when the highest-risk conditions are seen, identification of this rare disease may improve. **(Endocr Pract. 2016;22:000-000)**

Abbreviations:
CD = Cushing disease; **CPT** = Current Procedural Terminology; **CS** = Cushing syndrome; **EMR** = electronic medical record; **ICD-9-CM** = International Classification of Diseases, Ninth Revision, Clinical Modification; **ID** = identification; **RR** = relative risk

INTRODUCTION

Cushing syndrome (CS) is a rare endocrine disorder resulting from excess corticosteroid exposure. International population-based studies have estimated the annual incidence of endogenous CS at 1.2 to 2.4 per million people (1-3), and a U.S.-based study estimated 8 per million people (4). The majority of CS cases result from Cushing disease (CD) or excessive adrenocorticotropic hormone production from a pituitary adenoma (5). Uncontrolled CD is associated with substantial morbidity including hypertension, glucose intolerance, increased cardiovascular risk, osteopenia and fractures, and nephrolithiasis (5). Patients with CD have more than double the mortality risk of patients with nonfunctioning pituitary tumors (2,6), and those with persistent hypercortisolism fare even worse, with a standardized mortality ratio of approximately 4.0 compared to both the general population and patients with other pituitary adenomas (7). Therefore, timely and adequate treatment is crucial.

Submitted for publication July 7, 2015
Accepted for publication December 20, 2015
From the [1]Partnership for Health Analytic Research, Beverly Hills, California; [2]Novartis Pharmaceuticals Corporation, East Hanover, New Jersey.
Address correspondence to Dr. Dasha Cherepanov; Director, Outcomes Research; Partnership for Health Analytic Research, LLC
280 S. Beverly Drive, Suite 404, Beverly Hills, CA 90212.
E-mail: dasha@pharllc.com
Published as a Rapid Electronic Article in Press at http://www.endocrinepractice.org on January 20, 2016. DOI:10.4158/EP15914.OR
To purchase reprints of this article, please visit: www.aace.com/reprints.

CD is formally diagnosed through a combination of biochemical analysis and imaging studies. However, diagnosing CD in the first place is a considerable challenge: the disease is characterized by a wide range of signs, symptoms, and associated comorbidities that overlap those of many other more common conditions, for example metabolic syndrome (8). In certain high-risk populations, the prevalence of undiagnosed CD ranges from 0.5 to 1% in hypertensive patients to 10.8% in older patients with osteoporosis and vertebral fractures (9). Delays in diagnosis for a median of 2 years have been observed (10). These challenges highlight the need to increase awareness of CD and its wide range of signs, symptoms, and associated comorbidities among all types of clinicians seen by CD patients, including primary care physicians, to promote earlier recognition of symptoms and shorten the time to correct diagnosis.

Due to the overlap with other conditions, screening with individual CD-associated signs and symptoms may lead to an excess number of false positives. Identification of uncommon or unusual combinations of signs and symptoms, or "signals," may be more helpful in facilitating the diagnosis, but the infrequency of CD means that encountering unusual combinations of symptoms is an unpredictable process and may easily be overlooked by providers as having any clinical significance. This may be especially true as CD patients are seen by a variety of providers and specialists (11), all of whom may treat only certain aspects of the disease and therefore fail to observe broader patterns. Prior research suggests that compared with endocrinologists, non-endocrinologists have more difficulty diagnosing CS (10,12).

A potential solution to this challenge may be found using data mining, the process of discovering difficult-to-detect patterns and combinations in existing data sets. For example, researchers have identified unexpected drug combinations (i.e., signals) associated with conditions such as hyperglycemia, which can be confirmed using more traditional epidemiological and animal-based research (13). Data mining has been applied to databases dedicated to identifying such signals (e.g., the U.S. Food and Drug Administration Adverse Event Reporting System) and healthcare administrative databases (14,15). Data mining has also been used to analyze rare conditions, such as Charcot foot, a rare and serious diabetes complication (16) to identify multiple new or poorly described sign and symptom associations that were not initially intuitive.

In this paper, we describe a similar approach to achieve the same goals for CD. We analyzed a large U.S. healthcare claims database using a data-mining approach, with the goal of identifying combinations of symptoms and comorbidities observed more frequently in CD than in non-CD patients, particularly unusual combinations of otherwise common clinical symptoms or conditions. We focused on combinations that were common enough in the general population to be less likely to be overlooked, as a first step towards identifying markers that can serve as initial aids to diagnosis of CD.

METHODS

Data Source

This was a retrospective matched case-control study designed to compare the risks of individual conditions, as well as of combinations of conditions among patients with CD versus matched controls. We used the Truven Health Analytics MarketScan Database, a Health Insurance Portability and Accountability Act-compliant administrative claims database (17). The database includes health insurance claims from large employers and health plans across the U.S. and contains de-identified adjudicated pharmacy and medical claims submitted for payment by providers, healthcare facilities, and pharmacies. Claims include information on each physician visit, medical procedure, hospitalization, drugs dispensed, dates of service or prescription, number of days of medication supplied, and tests performed. Member enrollment and benefit information as well as limited patient, provider, and hospital demographic information are also available.

Patient Selection Criteria

Our study included patients with claims for CD in a 5-year identification (ID) period (January 1, 2008 to December 31, 2012) and matched patients without CD in the same period. For CD patients, the index date was January 1 of the year of the first CS diagnosis observed in the ID period, and the year following the index date was the measurement year. Matched controls had the same index date and measurement year.

There is no specific International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) diagnosis code for CD. Based on published literature (4,18), we used the following algorithm to identify eligible CD patients (appendices available on request): (1) at least 1 medical claim for CS (ICD-9-CM diagnosis code 255.0); (2) at least 1 medical claim for pituitary neoplasm (ICD-9-CM 227.3 or 237.0), miscellaneous pituitary disorders and syndromes (ICD-9-CM 253.1, 253.4, 253.8, or 253.9), hypophysectomy (ICD-9-CM procedure codes 07.61-07.65, 07.68 or 07.69; Current Procedural Terminology [CPT] codes 61546, 61548, or 62165), cranial stereotactic radiosurgery (CPT 61796-61799, 77371, and 77372), or bilateral inferior petrosal sinus sampling (CPT 36012 or 75860) with cortisol or adrenocorticotropic hormone sampling (CPT 82530, 82533, or 82924); and (3) continuous enrollment in the measurement year.

Non-CD patients were selected from a random 5% sample among all patients in the database. Non-CD patients had no medical claim with CS as 1 of the listed diagnoses in the ID period and were continuously enrolled for at least 1 calendar year in the ID period. For non-CD patients who

were continuously enrolled in more than 1 calendar year, we randomly picked 1 calendar year as their measurement year. For each CD patient, 2 non-CD patients with the same age, sex, and region in the same measurement year were randomly selected into the final study cohort. We randomly and equally divided CD patients into either the development or validation dataset, and the matched controls were then assigned to the development or validation dataset accordingly.

Study Measures and Statistical Analysis

A literature search was performed to identify clinical conditions linked with CD. The resultant list was reviewed by experts and modified appropriately. Ultimately, a group of 47 conditions was identified that included signs, symptoms, and comorbidities associated with CD. Only conditions identifiable using ICD-9-CM codes (appendices available on request) were eligible, as insurance claims do not contain other diagnosis data. In the development dataset, we reviewed all claims to find evidence of any of these 47 conditions and calculated their rates. We divided the rate of each condition in the CD group by the rate in the non-CD group to calculate the relative risk (RR) of each condition. Our aim was to focus on conditions common enough in the CD population that their identification could be an aid to early diagnosis. The list was reduced by examining the rate of each condition in the CD group. The initial threshold for retention was set at ≥5%. In order to include less common conditions without creating an unmanageably large list, we included conditions that were present at a ≥1% rate if the RR of the condition was ≥15 (e.g., 15 times more common in the CD group). A final review of all RRs with respect to specified thresholds by an expert clinician (W.H.L.) resulted in a final list of 10 conditions. The RR of these 10 conditions was then recalculated in the validation data set. The list of all 990 2-way combinations of the 47 conditions was similarly culled, using rate, RR, and expert consultation, to a final list of 24 pairs of conditions. The RRs were calculated for these 24 pairs and confirmed in the validation data set. We reviewed all claims in the measurement year for patient demographics: age, sex, and geographic region.

All data transformations and statistical analyses were performed using SAS version 9.4 (SAS Institute, Cary, NC). Descriptive statistics including means, SD, medians, and percentages, were reported for all measures whenever applicable.

RESULTS

The claims database yielded 3,750 patients who met the ICD-9-CM algorithm for CD. These 3,750 were matched 1:2 with non-CD patients, matched exactly on age, sex, region, and year of identification, yielding 7,500 matched non-CD controls. All 11,250 patients were included in the study, divided equally between the development and validation datasets (Fig. 1). For each dataset of 5,625 patients, there were 1,875 CD and 3,750 non-CD patients. The development and validation dataset populations were similar to each other in mean age, sex distribution, and U.S. region distribution (Table 1).

The frequencies of the 47 conditions in the CD population and their RRs varied widely. The frequency of hypertension was 43.5% (RR 2.5), while for abnormal genital virilization it was 0.053% (RR infinite). After applying our rules and expert review, the final list of individual conditions had RR varying from 27.8 for hirsutism to 5.1 for female balding in the development study. In the validation study, the RR of localized adiposity was infinite (i.e., localized adiposity was not identified in any patients in the non-CD group) versus 18 in the development data. Otherwise, the general order of RR was preserved. An RR ≥10.0 in the validation study was observed for 5 conditions: localized adiposity, hirsutism, facial plethora, polycystic ovary syndrome, and abnormal weight gain (Table 2).

The various 2-way combinations of conditions were similarly examined, with a final list of 24 clinically relevant combinations subjected to validation (Table 3). The frequency of these combinations varied from 14.3% (serious infections/hypertension) and 13.1% (serious infections/psychiatric disorders) to under 1% (uncontrolled type 2 diabetes/premature menopause and metabolic syndrome/fracture). The RR was infinite for 4 pairs of conditions (i.e., not observed in the non-CD group): hypertension/hirsutism, serious infection/adrenal mass, type 2 diabetes mellitus/hirsutism, and uncontrolled type 2 diabetes/premature menopause. Three other combinations had an RR >100: weakness or fatigue/hirsutism, hyperlipidemia/adrenal mass, and type 2 diabetes mellitus/adrenal mass.

DISCUSSION

Patients with CD are up to 60 times more likely to have certain conditions such as hirsutism coded in their insurance claims than age-, sex-, and region-matched controls. They are more than 100 times as likely to have certain combinations of conditions such as serious infection/adrenal mass. Patients with CD often suffer through years of missed opportunities to diagnose and treat their condition. We used data mining to identify, out of the thousands of possible combinations, those combinations of ICD-9-CM diagnosis codes that are highly associated with CD. Our results cannot yet be used to screen large populations of patients for CD since we used insurance claims with minimal clinical information for our analysis. However, these findings may be the first step in a data-driven process that will yield more precise methods for earlier identification of CD.

We demonstrated that the risk of having any 1 of 10 key conditions selected by data mining and expert endocri-

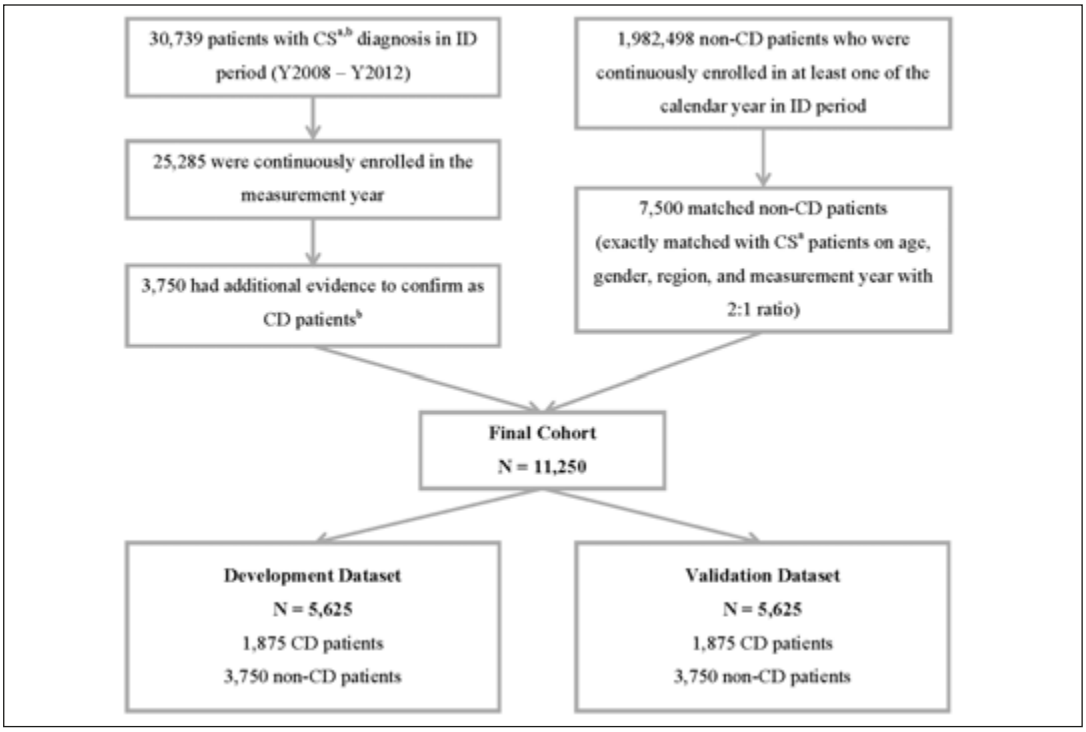

Fig. 1. Patient selection flowchart. *CD* = Cushing disease; *CS* = Cushing syndrome; *ICD-9-CM* = International Classification of Diseases, Ninth Revision, Clinical Modification; *ID* = identification.
[a]There is no specific ICD-9-CM code for CD.
[b]The CD patient identification algorithm used to identify the analytic cohort is was as follows: (*1*) ≥1 claim for CS (ICD-9-CM code 255.0); (*2*) ≥1 claim for pituitary neoplasm (ICD-9-CM 227.3 or 237.0), miscellaneous pituitary disorders and syndromes (ICD-9-CM 253.1, 253.4, 253.8, or 253.9), hypophysectomy (ICD-9-CM procedure codes 07.61-07.65 or 07.68; CPT codes 61546, 61548, or 62165), cranial stereotactic radiosurgery (CPT 61796-61799, 77371, and 77372), or bilateral inferior petrosal sinus sampling (CPT 36012 or 75860) with cortisol or adrenocorticotropic hormone sampling (CPT 82530, 82533, or 82924); and (*3*) continuous enrollment in the measurement year.

Table 1
Patient Demographics

CD n = 1,875		**Development dataset**		**Validation dataset**	
		Non-CD n = 3,750	**CD n = 1,875**	**Non-CD n = 3,750**	
Age, years	Mean	41.1	41.1	41.0	41.0
	(SD)	(13.0)	(13.0)	(13.1)	(13.1)
Female	n	1,440	2,880	1,433	2,866
	(%)	(76.8)	(76.8)	(76.4)	(76.4)
Region		351	702	340	680
North central	n				
	(%)	(18.7)	(18.7)	(18.1)	(18.1)
Northeast	n	456	912	486	972
	(%)	(24.3)	(24.3)	(25.9)	(25.9)
South	n	750	1,500	731	1,462
	(%)	(40.0)	(40.0)	(39.0)	(39.0)
West	n	318	636	318	636
	(%)	(17.0)	(17.0)	(17.0)	(17.0)

Abbreviation: CD = Cushing disease.

Table 2
RR of Selected Individual Conditions in CD Cases Versus Controls Based on Clinical Content Expert Opinion

Condition	Development dataset	Validation dataset
Localized adiposity	18.0	∞
Hirsutism	27.8	61.0
Facial plethora	15.0	21.0
Polycystic ovary syndrome	11.7	14.8
Abnormal weight gain	9.4	11.2
Hypokalemia	10.2	9.3
Deep venous thrombosis	10.8	7.5
Muscle weakness	8.0	7.3
Female balding	5.1	7.0
Osteoporosis	6.2	5.3

The 10 conditions in this table were chosen as follows: we initially selected for analysis 47 conditions defined by literature search and experts as being associated with CD in order to focus on conditions common enough in the CD population that their identification could be an aid to early diagnosis. We reduced the list by examining the frequency of each condition, setting an initial threshold for retention at ≥5%. In order to include less common conditions without creating an unmanageably large list, we included conditions that were present at ≥1% frequency if the RR of the condition was ≥15 (e.g., 15 times more common in the CD group). Iterative review by an expert clinician (W.H.L.) yielded further changes and resulted in a final list of 10 conditions, whose RRs were then recalculated in the validation data set. These results indicate that CD patients have a 5-fold or higher risk of having any 1 of these 10 conditions compared with non-CD patients.
Abbreviations: CD = Cushing disease; RR = relative risk.

nologist opinion is at least 5 times greater in CD patients compared to non-CD patients. Several individual conditions are predictably prevalent in CD. However, several of these are not specific to CD (e.g., hypertension) or very uncommon (e.g., adrenal mass) in the general population. Some combinations have been previously reported in the literature as case studies (e.g., psychiatric disorder/hirsuitism (11), hypertension/hirsuitism(18)), and others have been proposed to be more clinically suggestive of CD, such as osteoporosis-obesity (11,18-21). Findling and Raff previously described a list of signs, symptoms, and comorbidities that should initiate a biochemical evaluation for possible CS (25). Our list in Table 2 highly corresponds with this prior published list, except for deep venous thrombosis and female balding. The evidence of high RR of conditions in CD patients versus non-CD controls in our study, ranging from 61.0 for hirsutism to 5.3 for osteoporosis (and infinite RR for localized adiposity), highlights the importance of considering these conditions among standard screening criteria for possible CS. The value of our findings, resulting from an analysis of a large database of millions of insured subjects, is underscored by a consistent message resulting from a synthesis of clinical experience and published evidence from experts in the field.

An advantage of data mining over the clinically driven examination of combinations of conditions is the possibility of finding unexpected combinations. So while the previously proposed combination of osteoporosis/obesity is associated with an RR of 52 in the CD versus non-CD population, the combination of fatigue/hirsutism has an RR of 128. Advances in computing power have democratized the ability to mine large datasets to the point that such exercises using ICD-9-CM codes are nearly trivial from a computing standpoint. Unfortunately, ICD-9-CM codes themselves do not provide sufficiently rich data to do more than point the way for more testing. All the high-risk combinations we identified require further exploration. Two avenues for research appear particularly intriguing. First, confirmation of our findings in a medical record-based sample (e.g., a disease registry) could be undertaken. Such studies would be able to confirm the frequency of some newly identified combinations and establish that the diseases or conditions we identified are truly present (e.g., not a result of miscoding). They would not be able to estimate RR as registries do not typically contain disease-free controls. Second, data mining with a richer, more clinically detailed dataset could be undertaken. Both electronic medical record (EMR) and, to a lesser extent, ICD-10 data have

Table 3
RR in Patients with Versus without CD of Dyads of Conditions Selected by Clinical Content Expert Opinion

			Development dataset	Validation dataset				
				CD		Non-CD		
Order	Condition 1	Condition 2	RR	Frequency	(%)	Frequency	(%)	RR
1	Hypertension	Hirsutism	70.0	49	(2.613)	0	(0.000)	∞
2	Serious infections	Adrenal mass	∞	42	(2.240)	0	(0.000)	∞
3	Type 2 diabetes	Hirsutism	62.0	30	(1.600)	0	(0.000)	∞
4	Uncontrolled type 2 diabetes	Premature menopause	44.0	15	(0.800)	0	(0.000)	∞
5	Weakness/fatigue	Hirsutism	35.0	64	(3.413)	1	(0.027)	128.0
6	Hyperlipidemia	Adrenal mass	∞	56	(2.987)	1	(0.027)	112.0
7	Type 2 diabetes	Adrenal mass	∞	52	(2.773)	1	(0.027)	104.0
8	Psychiatric disorders	Hirsutism	43.0	49	(2.613)	1	(0.027)	98.0
9	Serious infections	Hirsutism	∞	44	(2.347)	1	(0.027)	88.0
10	Sleep disorders	Adrenal mass	∞	35	(1.867)	1	(0.027)	70.0
11	Uncontrolled Type 2 diabetes	Hypokalemia	15.0	32	(1.707)	1	(0.027)	64.0
12	Psychiatric disorders	Adrenal mass	∞	63	(3.360)	2	(0.053)	63.0
13	Weakness/fatigue	Adrenal mass	∞	56	(2.987)	2	(0.053)	56.0
14	Obesity	Osteoporosis	10.0	26	(1.387)	1	(0.027)	52.0
15	Metabolic syndrome/ impaired glucose tolerance/pre-diabetes	Vertebral, long bone, rib, pelvic and foot fracture	38.0	18	(0.960)	1	(0.027)	36.0
16	Weakness/fatigue	Female balding	15.3	30	(1.600)	3	(0.080)	20.0
17	Type 2 diabetes	Weakness/ fatigue	8.9	184	(9.813)	22	(0.587)	16.7
18	Type 2 diabetes	Hypokalemia	19.3	56	(2.987)	7	(0.187)	16.0
19	Obesity	Weakness/ fatigue	11.7	150	(8.000)	22	(0.587)	13.6
20	Hypertension	Osteoporosis	15.6	95	(5.067)	19	(0.507)	10.0
21	Type 2 diabetes	Premature menopause	31.3	35	(1.867)	7	(0.187)	10.0
22	Osteoporosis	Serious infections	15.1	66	(3.520)	16	(0.427)	8.3
23	Hypertension	Serious infections	4.6	269	(14.347)	106	(2.827)	5.1
24	Psychiatric disorders	Serious infections	4.5	245	(13.067)	119	(3.173)	4.1

Abbreviations: CD = Cushing disease; RR = relative risk.
The 24 pairs of conditions in this table were chosen as follows: we initially selected for analysis 47 conditions defined by literature search and experts as being associated with CD, in order to focus on conditions common enough in the CD population that their identification could be an aid to early diagnosis. After generating all possible 2-way combinations of these conditions, we chose for retention those that were either present at ≥5% frequency among CD patients or present at ≥1% frequency if the RR of the condition was ≥15. Using these frequency and RR data along with expert consultation, we narrowed the list down to 24 pairs of conditions. RR was calculated for these pairs and confirmed in the validation data set. These results indicate that CD patients have a 4-fold or higher risk of having any one of these 24 pairs of conditions compared with non-CD patients.

the potential to provide enough clinical detail that an even greater number of possible combinations could be mined—many millions in the case of EMRs (22). The rarity of CD makes use of either of these types of databases challenging at present. We used a database with an underlying sample size in the tens of millions, and it may be several years before EMR-based data sets of sufficient size are available.

Limitations

Our study has several limitations. This claims database, like most commercial claims databases, lacked laboratory test values, so we were not able to use biochemical test results to distinguish different forms of CS. The choice of conditions and combinations relied in part on expert opinion. Data mining does not require such an approach and can operate even in the absence of a priori hypotheses. We are considering expanding our analysis to include all ICD-9-CM codes to test whether even more unexpected combinations arise. As previously noted, insurance claims and associated ICD-9-CM codes lack clinical detail. Many conditions may be undercoded including both common ones (e.g., obesity) and uncommon ones. This problem may be particularly acute if the condition is peripheral to the reason the patient is seeking treatment (e.g., hirsutism in a patient treated for diabetes) (23). Inconsistency in reporting/coding over time and across health insurance plans and healthcare facilities can also result in unreliable data. Our data set covered a limited period of time, and some conditions may be coded once and not again. We would have missed those codes if they fell outside our study period. Generalizability to the rest of the population may be limited due to lack of inclusion of patients without commercial insurance (e.g., uninsured patients); those with insurance coverage may be healthier or more affluent and employed (24). Finally, our results of combinations of symptoms and comorbidities observed more frequently in CD than in non-CD patients should be useful in clinical practice but should be considered with caution given the rarity of this disease; even highly specific tests yield frequent false positives if the disease being tested for is rare.

CONCLUSION

Our analysis of a large U.S. healthcare claims database demonstrated that the RR of having any 1 of 10 key conditions selected by expert opinion was at least 5 times greater in CD compared to non-CD. Construction of uncommon dyads resulted in further increases in RRs beyond single condition analyses; for example, osteoporosis alone had an RR of 5.3, which increased to 8.3 with serious infections and to 52.0 with obesity. Nearly all dyads selected by expert endocrinologist opinion had RRs at least 5 times or greater. If clinicians consider the diagnosis of CD when the highest-risk conditions are seen, identification of this rare disease may improve. These results may be useful in developing clinical decision aids to identify patients at the highest risk of CD. Further studies using nonadministrative databases should be conducted to validate this research.

ACKNOWLEDGMENT

The authors thank Gordon H. Sun, MD, MS for his contributions to study design and drafting of the initial manuscript. Funding for this study was provided by Novartis Pharmaceuticals Corporation.

DISCLOSURE

Maureen P. Neary and William H. Ludlam are employees of Novartis Pharmaceuticals Corporation. Michael S. Broder, Eunice Chang, and Dasha Cherepanov are employees of the Partnership for Health Analytic Research, LLC, a health services research company paid by Novartis to conduct this research.

REFERENCES

1. **Etxabe J, Vazquez JA.** Morbidity and mortality in Cushing's disease: an epidemiological approach. *Clin Endocrinol (Oxf)*. 1994;40:479-484.
2. **Lindholm J, Juul S, Jørgensen JO, et al.** Incidence and late prognosis of Cushing's syndrome: a population-based study. *J Clin Endocrinol Metab*. 2001;86:117-123.
3. **Steffensen C, Bak AM, Rubeck KZ, Jørgensen JO.** Epidemiology of Cushing's syndrome. *Neuroendocrinology*. 2010;92 Suppl 1:1-5.
4. **Broder MS, Neary MP, Chang E, Cherepanov D, Ludlam WH.** Incidence of Cushing's syndrome and Cushing's disease in commercially-insured patients <65 years old in the United States. *Pituitary*. 2015;18:283-289.
5. **Newell-Price J, Bertagna X, Grossman AB, Nieman LK.** Cushing's syndrome. *Lancet*. 2006;367:1605-1617.
6. **Dekkers OM, Biermasz NR, Pereira AM, et al.** Mortality in patients treated for Cushing's disease is increased, compared with patients treated for nonfunctioning pituitary macroadenoma. *J Clin Endocrinol Metab*. 2007;92:976-981.
7. **Feelders RA, Hofland LJ.** Medical treatment of Cushing's disease. *J Clin Endocrinol Metab*. 2013;98:425-438.
8. **Feelders RA, Pulgar SJ, Kempel A, Pereira AM.** The burden of Cushing's disease: clinical and health-related quality of life aspects. *Eur J Endocrinol*. 2012;167:311-326.
9. **Nieman LK, Biller BM, Findling JW, et al.** The diagnosis of Cushing's syndrome: an Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab*. 2008;93:1526-1540.
10. **Valassi E, Santos A, Yaneva M, et al.** The European Registry on Cushing's syndrome: 2-year experience. Baseline demographic and clinical characteristics. *Eur J Endocrinol*. 2011;165:383-392.
11. **Tang A, O'Sullivan AJ, Diamond T, Gerard A, Campbell P.** Psychiatric symptoms as a clinical presentation of Cushing's syndrome. *Ann Gen Psychiatry*. 2013;12:23.
12. **Muller M, Mazzuco T, Martinie M, Bachelot I, Chabre O.** Diagnosis of Cushing's syndrome: a retrospective evaluation of clinical practice. *Eur J Intern Med*. 2006;17:334-338.

13. **Kuehn BM.** Scientists mine web search data to identify epidemics and adverse events. *JAMA*. 2013;309:1883-1884.
14. **Concaro S, Sacchi L, Cerra C, Fratino P, Bellazzi R.** Mining health care administrative data with temporal association rules on hybrid events. *Methods Inf Med*. 2011;50:166-179.
15. **Yoo I, Alafaireet P, Marinov M, et al.** Data mining in healthcare and biomedicine: a survey of the literature. *J Med Syst*. 2012;36:2431-2448.
16. **Munson ME, Wrobel JS, Holmes CM, Hanauer DA.** Data mining for identifying novel associations and temporal relationships with Charcot foot. *J Diabetes Res*. 2014;2014:214353.
17. **Truven Health Analytics.** MarketScan Research Databases. Available at: http://truvenhealth.com/portals/0/assets/ACRS_11223_0912_MarketScanResearch_SS_Web.pdf. Last update: 2012. Accessed on October 3, 2014.
18. **Nellesen D, Truong HL, Neary MP, Ludlam WH.** Economic burden of multiple chronic comorbidities associated with Cushing's disease, a rare endocrine disorder. *Expert Opin Orphan Drugs*. 2016;4:291-297.
19. **Abdel-Kader N, Cardiel MH, Navarro Compan V, Piedra Priego J, González A.** Cushing's disease as a cause of severe osteoporosis: a clinical challenge. *Reumatol Clin*. 2012;8:278-279.
20. **Guaraldi F, Salvatori R.** Cushing syndrome: maybe not so uncommon of an endocrine disease. *J Am Board Fam Med*. 2012;25:199-208.
21. **Yoshihara A, Okubo Y, Tanabe A, et al.** A juvenile case of Cushing's disease incidentally discovered with multiple bone fractures. *Intern Med*. 2007;46:583-587.
22. **Roque FS, Jensen PB, Schmock H, et al.** Using electronic patient records to discover disease correlations and stratify patient cohorts. *PLoS Comput Biol*. 2011;7:e1002141.
23. **Motheral B, Brooks J, Clark MA, et al.** A checklist for retrospective database studies. Report of the ISPOR Task Force on Retrospective Databases. *Value Health*. 2003;6:90-97.
24. **Schneeweiss S, Avorn J.** A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol*. 2005;58;323-337.
25. **Findling JW, Raff H.** Screening and diagnosis of Cushing's syndrome. *Endocrinol Metab Clin North Am*. 2005;34:385-402, ix-x. Review.