

Diagnosis Clusters: A New Approach for Reporting the Diagnostic Content of Family Practice Residents' Ambulatory Experiences

Ronald Schneeweiss, MB
Seattle, Washington

Effective documentation and monitoring of the ambulatory care experiences of residents have been elusive goals of family practice educators. This article describes a method of aggregating the diagnostic (morbidity) data recorded by the residents using diagnosis clusters. Reports organized by diagnosis clusters are useful and easily interpreted. They facilitate the monitoring and documentation of the diagnostic profile and the recording habits of each resident.

Family practice residents spend from one to five practice half-days practicing in the model teaching unit, depending on their level of training. Most graduates will spend 75 percent or more of their future practice time in ambulatory care¹; therefore, it is important to ensure that they obtain a sufficiently rich experience during their training to manage the majority of the problems they are likely to encounter in practice. Many facets of the resident's ambulatory experience warrant docu-

mentation and monitoring. In particular, the diagnostic content of the resident's practice is a useful indicator of the range of problems that he or she is managing. Effective monitoring demands readily interpretable aggregate reports that are easy for faculty and residents to comprehend. This article describes a method for documenting and monitoring the diagnostic content (morbidity profile) of a resident's practice during training that has been used in the University of Washington Affiliated Family Practice Network.

From the Department of Family Medicine, University of Washington, Seattle, Washington. Requests for reprints should be addressed to Dr. Ronald Schneeweiss, Department of Family Medicine, HQ-30, School of Medicine, University of Washington, Seattle, WA 98195.

University of Washington Network Information Management System (NIMS)

The NIMS system implemented in 1978 is based on information derived from the patient registra-

tion and encounter data in seven family practice residencies and one community practice. The profiles generated describe each resident's practice in terms of demographic, procedural, fiscal, and diagnostic variables. Aggregate reports have been developed incorporating peer-group and practice comparisons. A more detailed description of the current system has been published elsewhere.²

Problems Encountered in Reporting Morbidity Data

The many problems concerning the accuracy and reliability of the documented encounter data recorded by physicians have been well described by a number of authors³⁻⁷ and will not be dealt with in detail in this article. Much of the difficulty lies in the idiosyncratic labeling and coding habits of individual physicians and coders. The problem is compounded by adding the relative lack of training of the clerical staff, who often are responsible for the coding and related billing functions in ambulatory clinical settings.

An additional problem in describing the diagnostic content of practice is the need to reduce the large number of discrete diagnostic labels used in ambulatory family practice. The disease classifications used most commonly in the United States are the International Classification of Diseases (ICD-9-CM)⁸ and the International Classification of Health Problems in Primary Care (ICHPPC-2).⁹ ICD-9-CM has more than 10,000 diagnostic categories, and about 2,500 of these were used by family physicians to label the content of ambulatory practice in the National Ambulatory Medical Care Survey (NAMCS) 1980 and 1981.* ICHPPC-2 contains 367 categories, but even this number is overwhelming when attempting to describe and compare individual physicians' practices.

Previous attempts at displaying the data have included listing the discrete diagnoses in ranked order, which was the method adopted by the Virginia study on the content of family practice¹⁰ and also by NAMCS. Alternatively, diagnostic data have been grouped into the 17 ICD chapters, eg,

Chapter 5—"Diseases of the Circulatory System."^{11,12} This approach is too general for monitoring of resident experiences and blurs meaningful differences among providers. Other authors^{13,14} have selected certain conditions that they have arbitrarily chosen to monitor. This approach fails to reflect the overall spectrum of the content of that resident's practice; moreover, it does not highlight those areas where the resident is not getting sufficient exposure to common conditions occurring in that particular practice setting.

Diagnosis Clusters

Diagnosis clusters were developed as an instrument to facilitate the analysis of large data sets. The clusters grouped diagnoses that were considered to be clinically similar, thereby describing a physician's practice experience comprehensively while using a manageable number of categories. The diagnosis clusters were derived from the 1977 and 1978 NAMCS data sets for all specialties and included clinically meaningful groupings of conditions encountered in ambulatory practice. The clusters were shown to reduce the idiosyncratic labeling and coding preferences of providers, eg, low back pain (Table 1) as well as highlight conditions of low individual frequency but not high combined frequency (eg, fractures) (Table 2).

The 92 clusters included all discrete diagnoses or clusters of like diagnoses with a recorded frequency greater than 0.1 percent of all the diagnoses recorded. A detailed description of the clusters and the methodology used to derive them has been reported by Schneeweiss et al.¹⁵

When the NAMCS data for family physicians were analyzed, it was necessary to expand the clusters to 109 categories to encompass all diagnoses that had a frequency greater than 0.1 percent in the family practice office setting. The expansion of the original cluster roster is provided in Table 3.

Analysis of the NAMCS 1980 and 1981 tapes showed that the top 30 clusters accounted for 70 percent of all diagnoses recorded in the office setting by family physicians. The top 60 clusters accounted for 80 percent, and the roster of all 109 clusters accounted for 88 percent of the total recorded diagnoses (personal unpublished data).

The diagnosis clusters provide a means of ag-

*Unpublished data from the National Ambulatory Medical Care, 1980 and 1981, computer tape, available from the National Center for Health Statistics, Hyattsville, Maryland

Table 1. Low Back Pain: Example of Cluster That Reduces Idiosyncratic Labeling and Coding Preferences

ICDA-8	Cluster Content	Cluster Item Frequency (%)
<i>Low Back Pain Diseases and Syndromes (excluding acute strains)</i>		
353	Sciatica	0.12
713.1	Osteoarthritis of spine	0.19
717.0	Lumbago	0.01
725.1	Displacement of intervertebral disc lumbar and lumbosacral	0.16
725.9	Displacement of intervertebral disc NOS	0.14
728.7	Lumbalgia	0.15
728.9	Backache NOS	0.13
Aggregate Cluster Frequency		0.9

Table 2. Example of How Diagnosis Clusters Highlight Diagnoses With High Aggregate Frequency but Low Individual Frequency (NAMCS 1977-78 for all specialties*)

ICDA-8	Cluster Content	Cluster Item Frequency (%)
<i>All Fractures and Dislocations (excluding late effects):</i>		
807.0	Fracture of ribs	0.07
813.0	Fracture radius and ulna upper end, closed	0.11
813.4	Fracture of radius and ulna lower end, closed	0.12
816.0	Fracture of phalanges	0.21
824.0	Fracture of ankle	0.16
825.0	Fracture of metatarsals	0.12
800.0-828.0	All other fractures**	1.32
831.0	Dislocation of shoulder	0.03
836.0	Dislocation of knee—simple	0.12
830.0-839.0	All other dislocation†	0.08
Aggregate Cluster Frequency		1.78
*NAMCS, National Ambulatory Medical Care Survey, 1977-78		
**No discrete diagnosis in this group had a frequency greater than 0.09 percent		
†No discrete diagnosis in this group had a frequency greater than 0.04 percent		

gregating morbidity data in relatively few categories that, at the same time, retain clinical specificity and reflect the prevalence of diseases encountered in any given practice. In the family practice resi-

dency setting the use of diagnosis clusters permits meaningful comparisons between individual residents and peer groups using the overall practice experience as the norm.

Table 3. Specific Diagnoses Recorded by Family Physicians With a Frequency Greater Than 0.1 Percent Not Included in the Clusters Based on All Specialties (NAMCS 1977-78)¹⁵

Title	ICD-9-CM	ICHPPC-2	Percent of All Recorded Diagnoses
1. Viral infection unspecified	79.9	0799	0.43
2. Lipid disorders	272	272-	0.33
3. Edema	782.3	7823	0.27
4. Cough	786.2	7862	0.17
5. Hypertension	458	4580	0.17
6. Pleurisy	511.0-511.8	5110	0.17
7. Constipation	564.0	5640	0.16
8. Hypoglycemia	251.2	—	0.16
9. Lymphadenopathy	785.6	7856	0.15
10. Dyspnea	786.0, 786.1	7860	0.13
11. Skin tests	V72.7, V74.1	—	0.12
12. Venomous bites and stings	910-919 (all .4 and .5) 989.5	910	0.12
13. Nausea and vomiting	787.0	7870	0.12
14. Syncope	780.2	7802	0.11
15. Sleep disturbance	307.4	3074	0.11
16. Acute mastitis	675	676-	0.10
17. Ulcer of lower extremity (excluding decubitus)	707.1-.9	707-	0.10

Note: Since the above data were prepared, the North American Primary Care Research Group ad hoc Committee on Diagnosis Clusters has been developing a definitive cluster roster for family physicians, which will be made available at a later date

Morbidity Profiles Using Diagnosis Clusters

Figure 1 illustrates the morbidity report generated for each resident on a semiannual basis (the letters in parentheses are keyed to the figure). A recent survey of faculty and residents indicated that the report was found to be very useful by the majority of the respondents. The report indicates the reporting period and the physician as well as

the mean number of diagnoses per visit (a, b, c). In various studies the average number of diagnoses per visit for family physicians is remarkably consistent at 1.4 to 1.5 per visit. In this particular instance the provider, who was a second-year resident, recorded 1.5 diagnoses per visit. Recording only 1.0 to 1.2 diagnoses per visit would suggest that the resident is underrecording the problems seen.

The top 30 clusters (d) invariably account for

[a] JUL 1, 1982 TO JUN 30, 1983

FAMILY MEDICINE

[b]
Provider 2489

Percent of
Frequencies Practice

Number of Patient Visits 849 5.4
[c] Number of Diagnoses Recorded 1255 5.9
Mean # Diagnoses per Visit 1.5

[d]		Cluster Title (Cluster Number)	Frequencies		Percent of All Diagnoses		
Doctor Rank	Practice Rank		Number of Patients	Number of times Recorded	Cluster Percent	Cumulative Percent	Practice Cluster Percent
1	1	Obstetric Care (3)	[e] 41	99	[f] 7.9	7.9	9.2
2	3	Nonpsychotic Depression, Anxiety and Neuroses (5)	18	62	4.9	12.8	5.0
3	4	Acute Upper Respiratory Infection (2)	54	57	4.5	17.4	5.0
4	2	General Medical Examination (1)	39	53	4.2	21.6	8.3
5	6	Otitis Media - Acute and Chronic (11)	32	45	3.6	25.2	2.6
6	8	Contraception (40)	37	45	3.6	28.8	2.0
7	7	Vaginitis, Vulvitis, Cervicitis (21)	29	36	2.9	31.6	2.3
8	5	Hypertension (4)	14	34	2.7	34.3	4.6
9	9	Soft Tissue Injury (ex sprains) (6)	25	31	2.5	36.8	2.0
10	12	Menstrual Disorders (31)	23	30	2.4	39.2	1.6
11	25	+ Viral Infection Unspecified (93)	28	30	2.4	41.6	.7
12	19	Headaches (34)	21	26	2.1	43.7	1.1
13	11	Fibrositis, Myalgia, Arthralgia (33)	19	25	2.0	45.7	1.6
14	17	Dermatitis and Eczema (12)	21	24	1.9	47.6	1.3
15	10	Low Back Pain (27)	18	21	1.7	49.2	1.7
16	20	Nonfungal Skin Infections (26)	18	19	1.5	50.8	1.1
17	16	Abdominal Pain (64)	16	19	1.5	52.3	1.3
18	48	+ Constipation (99)	11	19	1.5	53.8	.4
19	21	Viral Warts (37)	12	18	1.4	55.2	1.0
20	15	Acute Lower Respiratory Infection (10)	16	17	1.4	56.6	1.4
21	32	Conjunctivitis, Keratitis (36)	13	16	1.3	57.8	.6
22	24	Acne and Diseases of Sweat, Sebaceous Glands (18)	11	15	1.2	59.0	.7
23	33	Peptic Diseases (29)	11	15	1.2	60.2	.6
24	28	Asthma (30)	8	14	1.1	61.4	.7
25	26	Hemorrhoids and Other Peri-Rectal Conditions (39)	11	14	1.1	62.5	.7
26	14	Urinary Tract Infection (19)	12	13	1.0	63.5	1.4
27	29	Benign and Unspecified Neoplasm (23)	10	12	1.0	64.5	.7
28	36	Debility and Undue Fatigue (89)	9	12	1.0	65.4	.5
29	55	+ Nausea and Vomiting (105)	11	12	1.0	66.4	.3
30	59	Vertiginous Syndromes (47)	8	11	.9	67.3	.3
[g]Practice top thirty clusters not included above.							
34	18	Acute Sprains and Strains (9)	9	9	.7	68.0	1.3
35	27	Degenerative Joint Disease (16)	3	9	.7	68.7	.7
36	30	Bursitis, Synovitis, Tenosynovitis (24)	6	9	.7	69.4	.6
39	23	Infectious Diarrhea and Gastroenteritis (25)	7	7	.6	70.0	.9
50	13	- Diabetes Mellitus (15)	2	3	.2	70.2	1.5
[h]Practice top thirty clusters not recorded by this provider.							
22	-	Congestive Heart Failure (63)	0	0	.0		.9

* Clusters for which the doctor's percent differs from the clinic percent by more than a factor of 3 are indicated by a + or -

Figure 1. Morbidity profile by diagnosis clusters

about 70 percent of all recorded diagnoses, which are compared with the practice mean for physician providers. Where the provider percent differs 3 times more or less from the practice, a + or - is indicated to the left of the cluster title. The number in parentheses refers to the cluster number in the index kept on file in each practice. The index includes the 92 clusters as reported in *Medical Care*¹⁵ plus the 17 categories listed in Table 3. The number of patients (e) and the frequency percent (f) for each cluster are also shown.

Reported below the top 30 clusters are the clusters (g) frequently encountered in the practice but not included in the provider's top 30. In this instance, the low frequency of diabetes mellitus is of note. These two segments provide a bird's-eye view of what the resident is actually seeing and recording in practice.

The most useful information in terms of facilitating faculty-resident dialogue is the section that documents the practice top 30 clusters *never* recorded by the resident (h)—in this instance, congestive heart failure. Clearly one would need to ensure that the residents obtained adequate experience in dealing with important common problems in ambulatory practice either by arranging for patients with those problems to be directed to their particular practice or by having the residents take a rotation in an appropriate specialty setting.

Comment

The diagnosis clusters aggregate the majority of problems that may be expected to be seen in family practice. One useful application is in documenting and monitoring the experience of family practice residents in the ambulatory setting. The morbidity recorded by the practice as a whole provides the reference against which each of the residents can be compared. The specific reporting format described in this article highlights those areas in which the resident may differ substantially from the practice norm. Since the report highlights deficiencies in the content of a given resident's practice, it is possible to correct them by directing patients with those conditions to that resident, as may occur when new patients join the practice or when senior residents graduate and their patients are reassigned. Differences in recorded morbidity

may be real or the result of poor recording habits, which is also a useful topic for discussion at the periodic faculty-resident evaluation meetings.

The one-page morbidity report presented here permits the faculty advisor and the resident to engage in a meaningful discussion based on the actual practice experience as documented by the resident. The report also provides a concise documentation of the diagnostic content of the resident's ambulatory experiences.

Acknowledgment

The members of the NIMS Committee contributed a helpful review of the cluster reporting format. The committee included Drs. Haigh Fox, Kenneth Gudgel, Rick Kirkwood, Mike Wanderer, and Mike Young, with Alberta Weinberg and Edward Gore (programmer).

References

1. Rosenblatt RA, Cherkin DC, Schneeweiss R, et al: The content of family practice: Current status and future trends. *J Fam Pract* 1982; 15:681-722
2. Schneeweiss R: Network information management system (NIMS). *Wash Acad Fam Pract J*, January 1981, p 11
3. Bentsen BG: The accuracy of recording patient problems in family practice. *J Med Educ* 1976; 51:311-316
4. Gehlbach SH: Comparing methods of data collection in an academic ambulatory practice. *J Med Educ* 1979; 54:730-732
5. Neuhaus E, Lyons TF, Payne BC: Problems of case finding and data collection in ambulatory care settings. *Am J Public Health* 1980; 70:282-283
6. Anderson JE: Reliability of morbidity data in family practice. *J Fam Pract* 1980; 4:677-683
7. Fortinsky RH, Gutman JD: A two-phase study of the reliability of computerized morbidity data. *J Fam Pract* 1981; 2:229-235
8. Ninth Revision, International Classification of Diseases, Clinical Modification (ICD-9-CM). Report of the Commission on Professional and Hospital Activities. Ann Arbor, University of Michigan, 1978
9. ICHPPC-2: International Classification of Health Problems in Primary Care. Second report of the Classification Committee of the World Organization of National Colleges, Academies, and Academic Associations of General Practice/Family Physicians. New York, Oxford University Press, 1979
10. Marsland DW, Wood M, Mayo F: The content of family practice. Part 1: Rank order of diagnoses by frequency. Part 2: Diagnoses by disease category and age/sex distribution. *J Fam Pract* 1976; 3:37-68
11. Boisseau V, Froom J: Practice profiles in evaluating the clinical experience of family medicine trainees. *J Fam Pract* 1978; 6:801-805
12. Rabin DL, Spector KK, Bush PJ: Ambulatory care in the community. *Pub Health Rep* 1980; 95:511-519
13. Tindall HL, Henderson RA, Cole AF: Evaluating family practice residents with a problem category index. *J Fam Pract* 1975; 2:353-357
14. Dickie GL, Newell JP, Bass MJ: An information system for family practice. Part 4: Encounter data and their uses. *J Fam Pract* 1976; 3:639-644
15. Schneeweiss R, Rosenblatt RA, Cherkin DC, et al: Diagnosis clusters: A new tool for analyzing the content of ambulatory medical care. *Med Care* 1983; 21:105