Table 1. Diagnoses Recorded by Nursing Staff From the Medical Record (n = 136)

Diagnosis	No.
Benign physical examination	41
No diagnosis stated	33
Cystic lesions	22
Cancer or rule out cancer or Paget's disease	18
Lump or fibroadenoma	11
Benign disease	5
"Dense tissue" or "thickening"	3
Rule out metastases	1
Pain	1
Nipple discharge	1

woman who claimed to do an examination that included both palpation and inspection found her disease by inspection; two women who claimed to find disease by inspection did not do palpation as part of their examination.

Comment

In this study, women claiming to do BSE were clearly not all doing the same examination. Further, it was found that no woman who included palpation found a disease by inspection. Simplifying BSE by limiting inspection to one arm position or by eliminating inspection may not reduce the efficacy of BSE in disease discovery. The present study supports this conclusion and the study findings are consistent with the evaluation of the physical examination by Mahoney and Csima.[4] A simplified examination may increase the confidence of women in BSE performance and thereby increase compliance.

Generalizability of the study is limited by the small sample size and by a possibly nonrepresentative patient population when compared with that of the average family physician's practice. Although further study is needed before definitive recommendations can be made, based on the results of Mahoney and Csima[4] and the present study, women should be taught to palpate the breasts while supine. Women should be informed that skin changes may be a sign of cancer and a physician should be contacted if any are noticed, but routine inspection is not so important as palpation. If routine inspection is done, it may be limited to one arm position (arms overhead).

References

1. Eggertsen SC, Bergman JJ: Breast self-examination: Historical perspective and current progress. J Fam Pract 16: 713, 1983
2. How to Examine Your Breasts, folder. New York, American Cancer Society, 1975
3. White LN, Cornelius JL, Judkins AF, et al: Screening of cancer by nurses. Cancer Nurs 1:15, 1978
4. Mahoney L, Csima A: Efficiency of palpation in clinical detection of breast cancer. Can Med Assoc J 127:729, 1982
5. Haagensen CD: Self-examination of the breasts. JAMA 149:356, 1952

Accuracy and Reliability of ICHPPC-2 Recording

Russell M. Boyle and Ronald Schneeweiss, MD

Richmond, Virginia, and Seattle, Washington

The *International Classification of Health Problems in Primary Care*, second edition (ICHPPC-2),[1] is widely used in recording morbidity from patient-physician encounters. Based on the Ninth Revision of the International Classification of Diseases (ICD-9), ICHPPC-2 is concise and flexible: it comprises 362 rubrics common in the ambulatory environment, and its optional hierarchical structure permits increased specificity in problem identification when desired.[2] This classi-

fication is currently in use in family practice residency programs[3-6] as well as community settings.[7]

Despite this prevalence, actual quality of morbidity data collected and reported remains unclear. The few relevant published papers concern the first edition of ICHPPC. Anderson[8] cited coding accuracy ranging from 92 percent to 97 percent, but the method of ascertaining these figures was not specified. Studies at the University of Massachusetts[3] and Brown University[9] compared contents of patients' medical records with automated morbidity data, and considerable discrepancies were detected. This paper presents measures of accuracy and reliability of ICHPPC-2 recording in two large data systems associated with family practice residency programs, the Virginia Family Practice Data System (VFPDS), Medical College of Virginia, and the Network Information Management System (NIMS), University of Washington.

Methods

Questions concerning morbidity recording procedures along with a listing of 18 written descriptions of problems were developed by the authors as the data collection form. This form included nine "common" problems seen frequently in family practice and nine problems not usually encountered or imprecisely identified. In both data systems, all participants at residency programs who routinely encode problems using ICHPPC-2 were asked to complete the form. Recording secretaries (office staff responsible for the central coding of problems into ICHPPC-2) and physicians (residents and faculty who do their own coding) were included. From a total of 28 people who entered coding data, there were 6 recording secretaries and 5 physicians from 4 VFPDS residency programs, and 11 recording secretaries and 6 physicians from 5 NIMS teaching practices. All forms were coded at the practices, reviewed by the authors for completeness, and sent to the Medical College of Virginia for analysis.

In this paper, *accuracy* is expressed as the percentage of problem descriptions coded correctly. The correct or most appropriate responses were determined by the authors through collaboration

with family physicians quite familiar with ICHPPC-2. *Reliability* is measured in percentage agreement with recorders in the same system. Accuracy and reliability findings for all respondents are presented in terms of 9 common, 9 atypical, and 18 total problems, while differences in accuracy between recording secretaries and physicians are based solely on total problems.

Results

Overall accuracy is displayed in Table 1: results are shown by type of problem description for each data system and both systems combined. Recorders fared much better with common as opposed to atypical problems, and an average of 70.4 percent of total problems was coded correctly by all respondents. Scores are strikingly comparable in both data systems. Table 2 illustrates that recording secretaries consistently coded problems more accurately than physicians.

Reliability results in Table 3 are again presented by type of problem for each data system. Expressed in average percentage agreement, reliability is also greater for common problems, while both systems show a low level of agreement for the atypical problems.

Discussion

Accuracy is best addressed in terms of the two types of problem descriptions. The mean of 90.9 percent for common problems (Table 1) reflects extremes of 44.4 percent and 100 percent. In fact, a rather impressive 5 of 11 VFPDS recorders and 9 of 17 NIMS recorders coded all common problems accurately. Problems coded correctly by all recorders in both data systems include "ear wax," "thyroid nodule," "abdominal pain," and "bursitis." The atypical group shows accuracy scores ranging from 22.2 percent to 77.8 percent, an expected finding because of these difficult problem descriptions. "Injection for allergy" and "chronic pain" were deemed "not codable" by the authors, and a considerable proportion of respondents (9 of 11 in VFPDS, 13 of 17 in NIMS) agreed.

Differences in accuracy between recording secretaries and physicians (Table 2) are not surprising. Some residents and faculty are especially

Table 1. Accuracy of All Respondents From the Virginia Family Practice Data System (VFPDS) and the Network Information Management System (NIMS)

Problem Descriptions	Average Percentage Correct		
	VFPDS (n = 11)	NIMS (n = 17)	VFPDS and NIMS Combined (n = 28)
Nine common problems*	91.9	90.2	90.9
Nine atypical problems**	47.4	51.6	49.9
All problems	69.7	70.9	70.4

*Nine common problems:
Ear wax, thyroid nodule, elevated blood pressure, abdominal pain, allergy-food, arteriosclerotic cardiovascular disease, benign prostatic hypertrophy, breast lump, bursitis
**Nine atypical problems:
Parosmia, Morton's neuroma, injection for allergy, cervical muscle strain, diabetic neuropathy, positive tuberculin, abdominal distress, contraception, chronic pain

Table 2. Accuracy of Recording Secretaries vs Physicians From the Virginia Family Practice Data System (VFPDS) and the Network Information Management System (NIMS)

System and Respondent	No.	All Problems Average Percentage Correct
VFPDS		
Recording secretaries	6	71.3
Physicians	5	67.8
NIMS		
Recording secretaries	11	75.7 ⎤ *
Physicians	6	62.1 ⎦
Both systems combined		
Recording secretaries	17	74.2 ⎤ **
Physicians	11	64.6 ⎦

*$P < .05$
**$P < .01$

proficient in their use of ICHPPC-2. Recording secretaries, however, tend to have more specialized training and experience in ICHPPC-2 morbidity recording. Thus, the implication for more accurate data from central as opposed to peripheral recording should be clear.

Table 3 indicates that recorders are much more likely to agree on common than atypical problems. The difference in reliability between these two groupings is quite substantial, reinforcing empiri-

Table 3. Reliability: Percentage Agreement With Recorders in Same System From the Virginia Family Practice Data System (VFPDS) and the Network Information Management System (NIMS)

Problem Descriptions	VFPDS Percentage Agreement (n = 55*)	NIMS Percentage Agreement (n = 136**)
Nine common problems	90.1	84.6
Nine atypical problems	62.0	54.1
All problems	76.4	70.2

$$^*n = \frac{11!}{2!(11-2)!} = 55$$

$$^{**}n = \frac{17!}{2!(17-2)!} = 136$$

cally that more disparate codes are used to record less frequent or vaguely identified problems. Agreement on all problems—roughly 76 percent in VFPDS and 70 percent in NIMS—connotes a reasonable level of uniformity in morbidity recording.

The similarity in the two data systems is noteworthy. Recording secretaries from NIMS are slightly more accurate than their peers from VFPDS for common, atypical, and total problems, while the combined VFPDS recorders agree more often than their NIMS counterparts: differences are indeed minimal. These findings demonstrate a degree of universality in ICHPPC-2 and suggest that comparisons of morbidity data between the two systems are at least feasible. Accuracy and reliability results from these two data systems must be viewed in a virtual vacuum, however, since comparative data from similar studies are not readily available.

This paper presents results of a pilot project designed as a precursor for a large-scale study. A comprehensive follow-up should thus produce listings of problems most likely to be coded accurately, those with more than one appropriate ICHPPC-2 code that may cause some divergence, and those (hopefully few) problem descriptions that will probably be coded incorrectly.

That it is important to measure accuracy and reliability in data systems utilizing ICHPPC-2 should be obvious. Only through these efforts can an estimate of the quality of recorded morbidity data be obtained, which is particularly important as the uses of patient-physician encounter data expand.

References

1. ICHPPC-2: International Classification of Health Problems in Primary Care. Report of the Classification Committee of the World Organization of National Colleges, Academies and Academic Associations of General Practitioners/Family Physicians. New York, Oxford University Press, 1979
2. Froom J: ICHPPC-2: An improved classification system for family practice. J Fam Pract 10:791, 1980
3. Rice CA, Godkin MA, Catlin RJO: Methodological, technical and ethical issues of a computerized data system. J Fam Pract 10:1061, 1980
4. Schneeweiss R: A network information management system (NIMS). Wash Acad Fam Phys J, January 1981, p 11
5. Boyle RM: Adding information and intelligence to a family practice data system. J Fam Pract 10:141, 1980
6. Green LA, Simmons RL, Reed FM, et al: A family medicine information system: The beginning of a network for practicing and resident family physicians. J Fam Pract 7:567, 1978
7. Kirkwood CR, Clure HR, Brodsky R, et al: The diagnostic content of family practice: 50 most common diseases in the WAMI community practices. J Fam Pract 15:485, 1982
8. Anderson JE: Reliability of morbidity data in family practice. J Fam Pract 10:677, 1980
9. Fortinsky RH, Gutman JD: A two-phase study of the reliability of computerized morbidity data. J Fam Pract 13:229, 1981