Evaluation of Abstracting:

Cancers Diagnosed in 1998

MCSS Quality Control Report 2004:1

Elaine N. Collins, M.A., R.H.I.A., C.T.R.

Jane E. Braun, M.S., C.T.R.

Wendy A. Mills, M.P.H.

February 2004

Minnesota Department of Health

Chronic Disease and Environmental Epidemiology

Minnesota Cancer Surveillance System

P. O. Box 9441

717 Delaware Street Southeast

Minneapolis, Minnesota 55440-9441

SUMMARY

The Minnesota Cancer Surveillance System (MCSS) performed a reabstracting study on a sample of 481 records from ten cancer registries for the primary cancer sites of head and neck, colon/rectum, female breast, cervix uteri, and bladder, and lymphomas diagnosed in 1998. Forty-six individual data items were reabstracted for each medical record, and 44 items were analyzed for this report. Data items were grouped into four categories. The combined (coding and computer transmission) weighted error rates ranged from 0.0% to 89.2% for individual data items, with many of the highest rates in the American Joint Committee on Cancer (AJCC) stage components calculated for the individual primary sites. Combined error rates were lower for demographic items (median 1.3%) than for treatment information (median 5.9%). Coding errors accounted for most discrepancies; software or transmission errors accounted for 17.3% of the total number of discrepancies in demographic data and 9.5% of the total number of discrepancies in treatment data. For coding only, overall agreement was 97.6% for demographic data items and 91.7% for variables pertaining to the first course of treatment. Coding agreement for cancer data items by primary site ranged from 85.3% for bladder to 93.4% for colon/rectum. Coding agreement for staging data items by primary site ranged from 51.5% for lymphoma to 88.5% for breast. Most discrepancies in demographic data related to race (assumed or not picked up from the record) and address (partially incorrect or not from the diagnosis date). Discrepancy rates over 10% in cancer fields were found in site for all the primary sites in the study; histology for breast, cervix, and lymphoma; grade for head and neck, bladder, and lymphoma; and diagnosis date for cervix. Discrepancy rates over 20% were found in tumor size for head and neck, colon/rectum, cervix, and bladder; nodes positive for lymphoma; nodes examined for head and neck and lymphoma; summary stage for lymphoma; certain TNM fields for all sites; AJCC clinical stage for all sites except breast; and AJCC pathologic stage for head and neck, cervix, and lymphoma. Many of the discrepancies in the staging fields are attributable to the grouping of stage variables between clinical versus pathologic basis for AJCC staging. Discrepancy rates over 10% in treatment fields were found in surgery, surgery date, reason no surgery, and treatment start date. Many of the discrepancies in the surgery fields are attributable to the change in coding standards for lymphoma biopsy procedures in 1998.

BACKGROUND AND PURPOSE

The MCSS began collecting information on cancer stage at diagnosis and first course of treatment for cancers diagnosed in 1995. In order to evaluate the quality of the stage and treatment data, as well as that of the demographic and cancer information, annual reabstracting studies were begun in 1997 with data from cases diagnosed in 1995. Formal reabstracting studies are commonly used to verify the accuracy of the data coded in the cancer registry against that contained in the medical record[1]. Accuracy is defined as the level of agreement between codes submitted to the central registry by the hospital registrars and coding assigned by an outside "expert" abstracter who codes the data without knowledge of the values previously assigned by the cancer registrar[2]. In accordance with North American Association of Central Cancer Registries (NAACCR) procedures, the MCSS elected to do comparative rather than blinded recoding. The purposes of the MCSS annual reabstracting studies are to: (1) estimate the overall and item-specific level of accuracy of the data submitted to the MCSS by hospital cancer registries, (2) identify systematic problems in collecting registry data which can be addressed through input to national standard-setting organizations, (3) identify areas where coding or interpretation of coding rules can be improved

through targeted training, (4) follow the estimated level of data accuracy over time, and (5) provide a mechanism for formal feedback to registrars.

METHODS

Cancers of the primary sites of head and neck, colon/rectum, female breast, cervix uteri, and bladder, and lymphomas diagnosed in 1998 were selected for the study. Thirty-three registry facilities were ranked by their total reported caseload (Minnesota residents only) for 1998; the two smallest-volume facilities were eliminated, and three strata were formed according to low, medium, and high reporting volume. A stratified systematic sample of facilities was chosen (every third facility), which included four facilities from the low volume stratum, four from the medium volume stratum, and two from the high volume stratum for the present study.

Up to ten records for each of five primary sites (excluding cervix uteri) were randomly selected from the reports submitted by each facility, for a total possible count of 500 records to be reabstracted for these cases. In the instances where a facility did not report at least ten cases for the primary site, all the eligible records were reabstracted. All cases of invasive cancer of the cervix uteri reported from the study facilities were included. The final sample size was 502 records.

Record reabstraction was conducted by one MCSS quality control staff person. Lists of study records were sent out prior to the study date, and the staff person spent approximately two days at each facility. For each record, 46 variables were reviewed and compared to a form containing the corresponding data from the MCSS database as submitted by the registry. New values were noted for each questioned field, and supporting documentation from the medical records was recorded. After record review, MCSS and registry staff discussed the data discrepancies found; depending on the facility, registry data were changed at the time of discussion, medical records were set aside for later review, or coding issues were noted.

All reabstracting forms were reviewed and assigned a reason for each discrepancy. During the first reabstracting study on 1995 data, it had been noted that data transmission formats and incomplete data transmission contributed to several discrepancies, particularly in the treatment fields. A two-tiered scheme was devised to allow for independently tabulating discrepancies caused by data coding and software issues, and this scheme has been maintained in all subsequent studies.

Data coding discrepancies were divided into seven major categories: missed data, updated information in the registry abstract not sent to the MCSS, coding errors, nonstandard registry coding practice, software-restricted coding, situations with unclear coding rules, and situations with conflicts between the reporting requirements of the MCSS and the American College of Surgeons. Discrepancies classified in the last two groups were not counted in the analysis. Unverifiable data that the registry accepted from another facility were also identified but accepted as given, unless contradicted in the facility or registry records.

Data were double-entered into a database with verification by MCSS data management staff. Discrepancies between the MCSS reabstracted data and the registries' coded data were enumerated, and percent disagreement was computed for each field. Total percent disagreement over the three strata was computed as weighted averages, with the proportions of cases within strata used as

weights. Analyses of demographic and treatment data were not stratified by anatomic site. For analyses of site-specific cancer and stage data, the proportion of cases within strata (weights) were computed using site-specific distributions of potential cases, assuming the original assignment of facilities to strata by volume was the same for all primary sites.

The number of records evaluated for each data item, and thus the denominator for each analyzed variable in the study, varied depending upon certain characteristics of the cases reviewed. In cases where the reported primary site was determined to be incorrectly coded to a study site, the records were evaluated only for demographic variables plus site, laterality, histology, and diagnosis date. In three cases the patient was determined to be a non-Minnesota resident, and the records were evaluated only for demographic variables. In terms of AJCC staging, the MCSS requires that registries report only one set of staging variables (clinical or pathologic) for each case; if the study record contained only clinical T, N, M, and stage group values, the pathologic T, N, M, and stage group fields were considered not applicable and not included in the denominator, and conversely for reported pathologic values and empty clinical fields. However, if the study record contained both clinical and pathologic AJCC staging values, both groups of variables were evaluated; and if the study record contained only one set of AJCC staging variables and the other set was considered to be better supported by the staging criteria documented in the medical record, both groups of variables were again evaluated. As discussed in a previous report, inpatient admission and discharge dates were reviewed during data collection but were eliminated from data analysis[3].

RESULTS

Twenty-one records were deleted from the original sample of 502: 20 records for which facility charts were unavailable, and one record from a pair of duplicate case submissions to the MCSS which had both been selected for the study. The final number of records analyzed was 481.

Table1 lists by data category (demographic and treatment variables for all records combined, cancer and staging variables by primary site) the percent agreement in coding for all variables by facility stratum, and total agreement rates (weighted by stratum) for software and coding. Tables 2 through 9 list, for each variable, the number and weighted percent of records with software problems or data coding discrepancies and the combined total number and weighted percent of records with either one or both types of discrepancies. The data coding fields are presented by facility stratum and total. Tables 2 and 9 present the demographic and treatment information, respectively. Tables 3 through 8 present cancer and staging information for the six primary sites.

Figure 1 presents agreement rates for data coding in the categories of demographic data and treatment data by facility stratum and overall. Figure 2 presents total agreement rates for data coding in the categories of cancer and staging data by primary site. Figures 3 and 4 present total numbers of data coding and software discrepancies for each study variable. Figure 5 presents the five categories of coding discrepancies as percentages of all coding discrepancies for demographic, cancer, staging, and treatment data. The discussion below focuses on specific coding issues from among these categories.

Demographic Fields

Demographic data variables included patient name, address, sex, date of birth, race, and social security number. The data coding agreement rate over all demographic fields was 97.6% (Table 1, Figure 1). The combined (coding and software) weighted percent of records containing discrepancies was 10.6% for address and 10.9% for race. Registrars assuming a patient's race when none was documented in the record accounted for 24 of the 47 coding discrepancies in this field. Race was missed in 17 cases, miscoded in four, and not updated to the MCSS in two cases.

Software problems with the address variable highlight issues in developing links between master patient and registry systems in a facility, whereby data are downloaded from a facility-wide system into a registry database. In eight cases, an apartment number was entered before a street number, and only the apartment number transmitted to the MCSS through the NAACCR reporting format. In five cases, the address (downloaded at the time of abstracting) was not the address of the patient at the time of diagnosis; in two of these five cases, the patient was not a Minnesota resident at diagnosis and the cases were not reportable to the MCSS. The coding problems with address reflect similar issues in manual data entry: in five cases, a building name or Post Office box number was entered before a street number and picked up in the NAACCR reporting layout; in nine cases, a later address was abstracted rather than the address at diagnosis, resulting in one non-Minnesota case being reported to the MCSS. The remaining coding errors in the address field reflect missing or partial addresses or incorrect spellings.

Cancer Fields

Cancer data fields were primary site, laterality, histology (including behavior), grade, and date of diagnosis. Agreement rates varied by primary site. The total weighted coding agreement rates by site were: head and neck, 90.1%; colon/rectum, 93.4%; breast, 90.3%; cervix uteri, 90.0%; bladder, 85.3%; and lymphoma, 88.5% (Table 1, Figure 2). The rate of discrepancies for individual fields also varied by primary site. Fields with coding discrepancy rates over 10% for individual sites were: head and neck—28.0% for site and 14.4% for grade; colon/rectum—15.5% for site; breast—20.4% for site and 14.9% for histology; cervix uteri—18.8% for site, 16.1% for histology, and 11.8% for diagnosis date; bladder—48.5% for site and 15.5% for grade; lymphoma—30.7% for site, 11.0% for histology, and 10.7% for grade (Tables 3-8). None of the cancer fields was affected by software problems.

Site was the cancer field with the most discrepancies noted. Of the 132 total coding discrepancies in site, 31 were major discrepancies resulting in a change of primary site for 11 cases each of colon/rectum cancer and lymphoma, and nine cases of head and neck cancer. Of the minor coding discrepancies in site, 60 involved records in which a general site code had been chosen when a specific subsite was documented (ten head and neck, two colon, five breast, ten cervix, 25 bladder, eight lymphoma). In 11 records, including eight lymphoma cases, multiple subsites of involvement were coded to a single subsite. In 30 cases, documentation supported coding of a different subsite or overlapping site code for the primary site.

Thirty-three of 46 coding discrepancies in histology were major discrepancies, resulting in a change in the first three digits of the histology code or a change in behavior between in situ and

invasive (four bladder cases). Two cases involved coding diagnoses that differed between an initial and a final or a revised report; one bladder histology which changed from invasive to in situ was not updated by the registry; in one case, an unconfirmed pathologic diagnosis was coded on the basis of clinical opinion. All other discrepancies (other than the in situ/invasive changes) involved not applying the best code to the stated diagnosis.

Coding discrepancies for laterality occurred in head and neck, breast, and lymphoma cases. Only one breast record involved a discrepancy between right and left, and one a discrepancy between unknown and known laterality. All discrepancies for head and neck cancer and lymphoma involved changes in site code between sites with and without laterality.

Twelve of 47 coding discrepancies within grade resulted from not picking up a grade from the pathology reports. Grade in three records was coded from a metastatic site; grade in six lymphoma records was coded from Working Formulation histologic designations rather than B-cell or T-cell classifications. In five records, grade coding varied from the ROADS terminology conversion table. Most of the remaining discrepancies resulted from coding grade incorrectly from the highest statement of grade given on single or multiple pathology reports.

Of the 20 coding discrepancies noted in the date of diagnosis, only three resulted in a date change greater than one month. Two of the outlying dates related to dates earlier than date of diagnosis (suspicious fine needle aspirate and negative biopsy recorded as positive), and one related to a date later than date of diagnosis (missed positive fine needle aspirate).

Staging Fields

Staging data variables were tumor size, nodes positive, nodes examined, metastases, summary stage, and American Joint Committee on Cancer (AJCC) clinical and pathologic T, N, M, and stage group. The data coding agreement rate varied by primary site: head and neck overall agreement was 76.3%; colon/rectum, 85.5%; female breast, 88.5%; cervix uteri, 74.4%; bladder, 82.4%; and lymphoma, 51.5% (Table 1, Figure 2). The highest combined (coding and software) discrepancy rates by site were: head and neck─37.7% for AJCC path group and 36.7% for pathologic T and N; colon/rectum─35.2% for AJCC clinical group, 24.7% for clinical N, and 22.2% for tumor size; female breast─26.9% for clinical T and 18.0% for pathologic T; cervix uteri─51.5% for clinical T, 44.5% for clinical N, and 43.5% for AJCC clinical group; bladder─35.8% for clinical T, 27.5% for AJCC clinical group, and 26.6% for pathologic T; lymphoma─ 89.2% for clinical M, and 77.9% for clinical T and clinical N (Tables 3-8). Software problems (across sites) affected the reporting of clinical M in seven records, AJCC clinical group in ten records, and AJCC pathologic group in two records.

Coding discrepancies were noted in tumor size in 114 of the records reviewed. Almost half of the discrepancies (in 49 records from all primary sites excluding lymphoma) occurred because tumor size was not coded when available. Size coded for something other than tumor dimension was the source of discrepancy in another 30 records, size was incorrectly coded in 15 records, and size was coded for lymphoma in nine records. Remaining discrepancies were due to miscellaneous causes.

Coding discrepancies in nodes positive and nodes examined were found in 67 and 80 records, respectively. Most discrepancies within the two fields resulted from: miscounting of nodes on the pathology report (four records for nodes positive and 17 records for nodes examined); coding nodes removed at surgery after neoadjuvant treatment (ten records for each of the two fields); using an incorrect code pairing, usually a transposition between the two fields (14 and 15 records, respectively); and using incorrect codes for lymphomas (34 and 33 records, respectively). Remaining discrepancies were due to miscellaneous causes.

Causes of coding discrepancies in the metastases fields included missing all or some involved sites (18 records), and coding uninvolved sites (six records). Other causes included coding progression of disease as metastatic at diagnosis, coding metastases for lymphomas, and miscoding involved sites.

Coding agreement rates for the three staging schemes varied by primary site (Tables 3-7). Agreement for summary stage was 96.2% for breast and 90.1% for bladder. Agreement was between 80% and 89% for summary stage for head and neck, colon/rectum, and cervix uteri; AJCC clinical stage for female breast; and AJCC pathologic stage for colon/rectum, breast, and bladder. Agreement was between 70% and 79% for summary stage for lymphoma; AJCC clinical stage for head and neck and bladder; and AJCC pathologic stage for cervix. Agreement was less than 70% for AJCC clinical stage for colon/rectum, cervix uteri, and lymphoma; and AJCC pathologic stage for head and neck and lymphoma. Agreement for the individual T, N, and M components of the AJCC stage group is also detailed in Tables 3-8.

A major source of AJCC staging discrepancies was clinical/pathologic staging conflicts. These conflicts included: the use of pathologic staging elements in clinical staging when clinical T and N were not available; assignment of pathologic staging in cases where the primary surgical procedure did not meet the pathologic criteria; assignment of clinical staging only in cases meeting the criteria for pathologic assignment; and assignment of pathologic staging only in cases with neoadjuvant treatment where clinical staging is preferred. Another source of discrepancy in staging was the use of an incorrect form where a case was not staged, such as 88 (not applicable) for 99 (unknown stage), considered a minor discrepancy. The very high discrepancy rates in the clinical and pathologic TNM fields for lymphoma are primarily due to a registry software program not allowing the appropriate recording of 88 in these fields as not applicable for lymphoma.

For AJCC clinical stage group, the number of discrepancies attributed to clinical/pathologic staging conflicts were: head and neck, seven; colon/rectum, six; breast, four; cervix uteri, six; bladder, five; and lymphoma, 12. This type of discrepancy accounted for 45.2% of all discrepancies in this field. For the other discrepancies, the numbers of major discrepancies (resulting in a change between stage groups) and minor discrepancies (resulting in a change within a stage group) for each primary site were: head and neck, six major and three minor; colon/rectum, five major and three minor; breast, six major; cervix uteri, four major; bladder eight major and six minor; and lymphoma, seven major and 16 minor.

Common reasons for AJCC clinical stage discrepancies across primary sites included incorrect assignment of T values and the exclusion of CT and MRI findings. Most of the discrepancies in head and neck cancers related to the selection of T and N values from multiple clinical reports

within the record. Discrepancies in lymphomas arose in applying the definition of multiple nodal regions for staging purposes and in determining Stage IV disease for extranodal lymphomas. The minor discrepancies for bladder resulted from the inappropriate coding of 0A versus 0IS for papillary versus nonpapillary urothelial tumors.The minor discrepancies in the lymphoma records resulted from the coding or non-coding of B symptoms.

For AJCC pathologic stage group, the number of discrepancies attributed to clinical/pathologic staging conflicts were: head and neck, nine; colon/rectum, five; breast, five; cervix uteri, six; bladder, nine; and lymphoma, 21. This type of discrepancy accounted for 50% of all discrepancies in this field. For the other discrepancies, the numbers of major and minor discrepancies for each primary site were: head and neck, five major and one minor; colon/rectum, four major; breast, five major and five minor; cervix uteri, two major and three minor; bladder, two major and three minor; and lymphoma, six major and nine minor.

Reasons for pathologic staging discrepancies varied with few common elements across sites and within sites for head and neck, colon, and bladder. Most discrepancies in breast cancer related to incorrect pathologic T value or incorrect coding of final stage group based on TNM values. Major discrepancies for lymphomas related to positive bone marrow biopsies that were not identified on abstracting, and incorrect staging for extranodal lymphomas with distant involvement. Minor discrepancies for lymphomas again related to coding or non-coding of B symptoms.

For summary stage, the numbers of major discrepancies (resulting in a change between stage groups) and minor discrepancies (resulting in a change within regional stage groups) for each primary site were: head and neck, eight major and four minor; colon/rectum, 14 major and four minor; breast, two major; cervix uteri, four major and two minor; bladder, nine major and one minor; and lymphoma, 19 major and three minor. Many of the discrepancies in head and neck, colon/rectum, bladder, and cervix uteri records related to coding descriptions of anatomic extent of disease that are contained within the summary stage scheme but are not specified within the AJCC scheme. For example, discrepancies occurred in cases of cancers of the cervix uteri extending to paracervical or ovarian tissue, specified in the summary stage scheme but not within AJCC. Discrepancies also occurred in head and neck cancers with extension to tissues not specifically mentioned within the site-specific summary stage guidelines, but where stage could be assigned within the general guidelines of the scheme. Many of the discrepancies in colon/rectum cancers related to identifying localized and regional stages from terms describing subserosal, serosal, and pericolonic involvement of tumor, and also identifying regional and distant disease from terms describing involvement of mesenteric and peritoneal tissues.

In the two breast summary stage discrepancies, one record omitted summary stage, and one record coded a localized tumor as regional. Discrepancies in summary stage for bladder related to not assigning an unknown stage code when appropriate, as well as incorrect coding of the extent of the primary tumor. Lymphoma cases with discrepancies in AJCC staging had discrepancies for similar reasons in summary stage. In addition, 12 lymphoma records had discrepancies because the codes for summary stage were not applied correctly, in terms of matching the AJCC stage value to the summary stage value (AJCC 1/SS 1, AJCC 2/SS 5, AJCC 3/SS 7, AJCC 4/SS7).

Treatment Fields

Treatment data included variables pertaining to type of therapy given and therapy start dates. The data coding agreement rate for treatment fields was 91.7% (Table 1, Figure 1). The highest combined error rates for coding and software occurred in treatment fields for surgery (28.7%), date of surgery (19.9%), reason no surgery (16.6%), and date treatment started (27.1%) (Table 9).

Software-associated problems primarily affected the surgery date (3.3%) and start treatment date (7.7%) fields. These problems related to issues in data from four software vendors. In two situations, dates were not picked up or reported; in two situations, incorrect dates were reported, including the translation of 0-filled dates to 9-filled by one reporting program.

Many of the coding discrepancies in the surgery field were related to the change in coding standards effective with the diagnosis date of January 1, 1998. A type of data discrepancy that was not counted as a software error for this study was the reporting of treatment in 22 records (from three facilities) coded using pre-98 surgery codes rather than the codes implemented for diagnosis date of 1998; of these 22 records, six were updated with 1998 surgery codes after the study sample was drawn. A related type of discrepancy counted among the coding errors was the reporting of 1998 surgery codes probably converted from pre-1998 codes in 13 records, with a resulting lack of specificity in the code when compared with the medical record. In an additional 20 cases, a less specific code was also applied to the surgical procedure, though conversion issues did not seem to apply to these cases. Thirty-seven coding discrepancies can be attributed to the change in coding lymphoma procedures with the 1998 instructions to code lymph node biopsies as a surgical procedure. Most of the remaining discrepancies resulted from not applying the best code to the described procedure.

Corresponding to the records with surgery discrepancies, 36 lymphoma cases also had discrepancies in surgery date. Twenty-five cases involved a conflict between incisional and excisional biopsy dates for other primary sites. In eight cases, a treatment date was missed or an updated value not transmitted to the MCSS. The discrepancies in the treatment start date field follow a similar pattern. Discrepancies in this field involved 30 lymphoma cases and 39 conflicts between excisional and incisional biopsy dates for other primary sites. The date was imprecise by one or two days in eight cases, and treatment was missed in seven cases.

Similar to the coding discrepancies in the surgery field, 34 discrepancies in the reason no surgery field can be attributed to the change in coding excisional node biopsies for lymphoma as treatment starting in 1998; five discrepancies also related to the relationship of this field to two additional surgery fields that were added in 1998 (though not reported in this study): scope of regional node surgery, and surgery of other regional and distant sites. Another 16 discrepancies are related to reporting a software-supplied default for the field versus registry coding. The remaining discrepancies related to values not coded or not updated to MCSS, or the misapplication of codes to record data.

Most discrepancies within the radiation treatment field were for treatment that was missed or not updated to the MCSS (ten of 17 records from four of six primary sites), or for subsequent treatment reported as first course. Eight of the discrepancies in chemotherapy resulted from coding treatment

for colon cancer with 5-Fluorouracil-Leucovorin as multiple-agent rather than single-agent therapy, while another eight discrepancies resulted from coding chemotherapy as not otherwise specified when more specific information was available. Chemotherapy was missed or not updated in another five cases, and subsequent therapy was coded as initial therapy in three records.

Most discrepancies in the hormone therapy field (26 of 40) resulted from not coding prednisone when given as part of the treatment regimen for lymphoma. Hormone therapy was missed or not updated to the MCSS in another 12 cases (nine breast and three lymphoma cancers). Discrepancies in immunotherapy involved missed or not updated treatment or subsequent treatment coded as first course (six of seven cases), while discrepancies in other treatment were divided between missed and miscoded data (three cases each).

DISCUSSION

Few discrepancies were found in demographic data. Problems with identifying and coding race have been noted in previous studies. This study also noted a specific problem with the identification of address at diagnosis, which may not be abstracted correctly if registrars download patient addresses into their registry databases at the time of abstracting and do not verify addresses with records from the time of diagnosis.

Cancer data items were generally in agreement with the information in the medical record. Most discrepant diagnosis dates were accurate to within one month of the actual date. Training may prove most beneficial in a basic review of site and histology coding. Grade continues to be a difficult field to code, and clarification of the standards for this field may be most helpful to achieve consistency in coding of this data item.

The distinction between clinical and pathologic staging had the greatest impact on staging for lymphoma and the least impact on staging for breast. The disagreement rates for AJCC clinical staging for colon/rectum and breast cancers were most affected by low numbers. Most registries provide only pathologic staging for these sites, and cases with valid pathologic staging and no clinical staging were not included in the denominators for AJCC clinical staging. Likewise, the disagreement rate for AJCC pathologic staging for bladder cancer was most affected by low numbers, as most bladder cancers are clinically staged. Specific areas in which training may improve data quality include the appropriate use of clinical information in clinical and pathologic staging, a review of summary stage for colon/rectum cancers and summary stage coding for lymphomas, and reinforcement of the use of the summary staging guide when abstracting cases.

The distinction between clinical and pathologic staging elements, and registrars' reliance on staging forms that physicians use for clinical purposes, create recognized problems in the precise coding of AJCC stage data. The change to the collaborative staging system for recording registry data in 2004 should resolve the data quality issue represented by the clinical/pathologic staging conflict, which has been illuminated by all of the MCSS reabstracting studies. The publication of the *AJCC Cancer Staging Manual Fifth Edition* for use with 1998 cases did not appear to have a discernible effect on staging quality.

The lymphoma cases illustrate two types of impact that a change in data standards can have on data quality. The *Registry Oncology and Data Standards* (ROADS) manual for coding was introduced in 1996 and revised with a new surgery coding schema in 1998. The very high rates of discrepancies in the AJCC TNM fields for this site are primarily attributable to a registry software program that was not revised to accommodate the 1996 ROADS standard for coding 88 in the TNM fields when not applicable. The high rates of discrepancies in the surgery fields are largely attributable to registrars not immediately adopting the 1998 revised coding standard for excisional biopsy procedures.

The treatment fields illustrate an additional type of impact on data quality flowing from the variable timeframe within which standards are implemented. Some records in this study may have been abstracted and reported before the registry's software was updated to accommodate the new surgery codes, and treatment data later added to the abstract were not updated to the MCSS. (This situation was considered an administrative problem for this study, and the unknown treatment data in such records were not counted as errors.) Other records in this study were abstracted under the old standards, and treatment data were later converted to the new standards and updated to the MCSS, resulting in a loss of specificity in the coding. The publication of implementation guidelines by the NAACCR for a similar change in data standards with 2003 cases may help to resolve the data quality issues represented by standards implementation schedules.

Comparing the results of this 1998 reabstracting study with prior data studies demonstrates that the impact of software problems on data quality has been stable for the last two studies. The overall software discrepancy rate was 0.9% for treatment fields in 1998, compared to 0.7% in 1997, 2.0% in 1996, and 15.6% in 1995. Coding agreement rates were comparable over time for demographic fields: 97.6% in 1998, 98.0% in 1997 and 1996, and 96.7% in 1995. Site-specific coding agreement rates in cancer fields were somewhat lower in 1998, ranging from 85.3% for bladder cancer to 93.4% for colon/rectum cancer, compared with 90.1% to 97.3% in 1997, 90.0% to 97.4% in 1996, and 87.0% to 96.9% in 1995. The range of site-specific coding agreement rates in staging fields was much wider in 1998, ranging from 51.5% for lymphoma to 88.5% for breast cancer, compared with 75.9% to 85.2% in 1997, 75.5% to 88.6% in 1996, and 80.7% to 91.3% in 1995.

Female breast has been included as a study site for all study years, so that comparisons in data quality for one site could be made across all facilities and over time. Coding agreement rates for breast cancer fields were 90.3% in 1998, 90.1% in 1997, 90.5% in 1996, and 92.2% in 1995. Coding agreement rates for the single summary stage field for breast were 96.2% in 1998, 92.5% in 1997, 94.3% in 1996, and 99.4% in 1995. Coding agreement rates for all breast staging fields were 88.5% in 1998, 83.6% in 1997, 88.6% in 1996, and 91.3% in 1995.

Comparative agreement rates for other 1998 study sites that have been in previous studies are as follows: head and neck─cancer 90.1% and staging 76.3% in 1998, cancer 93.4% and staging 82.1% in 1995; colon/rectum─cancer 93.4% and staging 85.5% in 1998, cancer 97.3% and staging 85.1% in 1997; bladder─cancer 85.3% and staging 82.4% in 1998, cancer 82.6% and staging 80.7% in 1995; lymphoma─cancer 88.5% and staging 51.5% in 1998, cancer 87.0% and staging 84.8% in 1995.

Data have not been formally analyzed by the strata of low, medium, and high volume facilities. The facilities were ranked on the volume of Minnesota cases reported for all sites in 1998. This ranking does not reflect total registry caseload for any non-Minnesota facilities near the border which report to the MCSS; the Minnesota caseload makes up a small percent of the total case volume at these facilities.

FUTURE PLANS

Individual reports from the reabstracting study will be prepared for each participating facility, focusing on areas of coding disagreement, providing comparisons to the study group as a whole and to facilities of similar size, and providing information to improve future data quality. The MCSS will continue to include breast cancer in all reabstracting studies, and will be revisiting facilities and primary sites on a three-year cycle; the MCSS plans to monitor trends in data quality for breast cancer specifically, and for other sites as comparative data become available. The MCSS will continue to sponsor training workshops focusing on data quality issues, bringing in national speakers and also developing presentations by MCSS staff. The MCSS will continue to encourage appropriate workshops hosted by the Minnesota Cancer Registrars Association (MCRA), and will continue to contribute articles on data quality issues to the newsletter published by the MCRA. The MCSS will continue to work closely with software vendors to assure that data can be abstracted according to current standards and are transmitted in the required reporting formats.

Through future reabstracting studies, the MCSS will continue to track the impact of changing data standards on data quality. New standards implemented since the collection of data for cases diagnosed in 1998 include the *International Classification of Diseases for Oncology Third Edition* and the *SEER Summary Staging Manual 2000*, both implemented for cases diagnosed in 2001, and the *Facility Oncology and Registry Data Standards* (FORDS), implemented for cases diagnosed in 2003. The collaborative staging system will be implemented for cases diagnosed in 2004. Ideally, the publication of new standards resolves data consistency issues previously identified, although each standard also brings new challenges in implementation and interpretation.

ACKNOWLEDGEMENT

Analytic programming support provided by Wendy Trute, M.P.H. and John Soler, M.P.H.

Table 1. Agreement in Data Categories

	Coding								Software	
	High		Medium		Low		Total *		Total *	
	No.	%	No.	%	No.	%	No.	%	No.	%
Demographic Fields										
N fields = 11										
Valid observations **	1,397		2,046		1,848		5,291		5,291	
Agreement	1,372	98.2	1,976	96.6	1,814	98.2	5,162	97.6	5,264	99.1
Treatment Fields										
N fields = 15										
Valid observations ***	1,860		2,775		2,490		7,125		7,125	
Agreement	1,715	92.2	2,553	92.0	2,219	89.1	6,487	91.7	7,058	99.1
Cancer Fields – Head/Neck										
N fields = 5										
Valid observations ***	100		155		115		370		370	
Agreement	90	90.0	143	92.3	98	85.2	331	90.1	370	100.0
Cancer Fields - Colon/rectum										
N fields = 5										
Valid observations ***	95		195		165		455		455	
Agreement	89	93.7	182	93.3	152	92.1	423	93.4	455	100.0
Cancer Fields - Breast										
N fields = 5										
Valid observations ***	95		155		195		445		445	
Agreement	85	89.5	143	92.3	173	88.7	401	90.3	445	100.0
Cancer Fields - Cervix										
N fields = 5										
Valid observations ***	145		50		54		249		249	
Agreement	129	89.0	46	92.0	48	88.9	223	90.0	249	100.0
Cancer Fields - Bladder										
N fields = 5										
Valid observations ***	95		185		135		415		415	
Agreement	80	84.2	157	84.9	122	90.4	359	85.3	415	100.0
Cancer Fields - Lymphoma										
N fields = 5										
Valid observations ***	94		185		174		453		453	
Agreement	84	89.4	163	88.1	150	86.2	397	88.5	453	100.0
Staging Fields – Head/Neck										
N fields = 13										
Valid observations ***	204		347		247		798		798	
Agreement	147	72.1	282	81.3	188	76.1	617	76.3	797	99.9

Table 1. Agreement in Data Categories (continued)

	Coding								Software	
	High		Medium		Low		Total *		Total *	
	No.	%	No.	%	No.	%	No.	%	No.	%
Staging Fields - Colon/rectum										
N fields = 13										
Valid observations ***	179		409		345		933		933	
Agreement	148	82.7	360	88.0	293	84.9	801	85.5	931	99.9
Staging Fields – Breast										
N fields = 13										
Valid observations***	191		319		423		933		933	
Agreement	174	91.1	279	87.5	341	80.6	794	88.5	928	99.8
Staging Fields - Cervix										
N fields = 13										
Valid observations ***	277		106		118		501		501	
Agreement	223	80.5	72	67.9	84	71.2	379	74.4	501	100.0
Staging Fields - Bladder										
N fields = 13										
Valid observations ***	171		433		291		895		895	
Agreement	139	81.3	358	82.7	228	78.4	725	82.4	895	100.0
Staging Fields - Lymphoma										
N fields = 13										
Valid observations ***	182		417		386		985		985	
Agreement	80	44.0	241	57.8	177	45.9	498	51.5	974	99.1

* Total percentages weighted by stratum size to reflect state total
** Valid observations = n fields x m records reabstracted
*** Valid observations = n fields reabstracted in m records (i.e., not all fields reabstracted in all records)

Table 2. Records with Discrepancies in Demographic Fields: All Sites Combined

	Coding Errors								Software Errors		Combined Errors	
	High		Medium		Low		Total *		Total *		Total *	
Field	**No.**	**%**	**No.**	**%**	**No.**	**%**	**No.**	**%**	**No.**	**%**	**No.**	**%**
Last Name	0	0.0	1	0.5	1	0.6	2	0.3	0	0.0	2	0.3
First Name	1	0.8	3	1.6	0	0.0	4	1.0	0	0.0	4	1.0
Middle Name	2	1.6	15	8.1	8	4.8	25	4.3	0	0.0	25	4.3
Address	4	3.1	17	9.1	7	4.2	28	5.4	13	5.2	41	10.6
City	0	0.0	7	3.8	1	0.6	8	1.4	5	2.0	13	3.4
State	0	0.0	1	0.5	0	0.0	1	0.2	2	0.8	3	1.0
Zip Code	2	1.6	6	3.2	2	1.2	10	2.1	2	0.8	12	2.9
Sex	0	0.0	0	0.0	1	0.6	1	0.1	0	0.0	1	0.1
Date of Birth	0	0.0	0	0.0	1	0.6	1	0.1	0	0.0	1	0.1
Race	16	12.6	18	9.7	13	7.7	47	10.9	0	0.0	47	10.9
SSN	0	0.0	2	1.1	0	0.0	2	0.4	5	0.9	7	1.3

* Total percentages weighted by stratum size to reflect state total.

Note: Data for the individual strata (high, medium, low) are not shown for the software and combined totals.

Table 3. Records with Discrepancies in Cancer and Staging Fields: Head and Neck

	Coding Errors								Software Errors		Combined Errors	
	High		Medium		Low		Total *		Total *		Total *	
Field	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Site	6	30.0	8	25.8	6	26.1	20	28.0	0	0.0	20	28.0
Histology	0	0.0	0	0.0	3	13.0	3	1.9	0	0.0	3	1.9
Laterality	0	0.0	1	3.2	4	17.4	5	3.6	0	0.0	5	3.6
Grade	4	20.0	2	6.5	3	13.0	9	14.4	0	0.0	9	14.4
Dxdate	0	0.0	1	3.2	1	4.3	2	1.7	0	0.0	2	1.7
Tumor Size	2	10.0	9	29.0	9	39.1	20	20.7	0	0.0	20	20.7
Nodes Positive	3	15.0	0	0.0	8	34.8	11	12.7	0	0.0	11	12.7
Nodes Examined	6	30.0	1	3.2	11	47.8	18	23.4	0	0.0	18	23.4
Distant Mets	1	5.0	1	3.2	2	8.7	4	4.9	0	0.0	4	4.9
Summary Stage	3	15.0	4	12.9	5	21.7	12	15.2	0	0.0	12	15.2
T Clinical	7	46.7	7	25.9	4	22.2	18	34.8	0	0.0	18	34.8
N Clinical	5	33.3	8	29.6	8	44.4	21	32.9	0	0.0	21	32.9
M Clinical	2	13.3	4	14.8	1	5.6	7	12.6	1	1.4	8	13.9
AJCC Clin Stage	4	26.7	9	33.3	3	16.7	16	27.3	0	0.0	16	27.3
T Pathologic	6	54.5	6	28.6	2	13.3	14	36.7	0	0.0	14	36.7
N Pathologic	6	54.5	6	28.6	2	13.3	14	36.7	0	0.0	14	36.7
M Pathologic	6	54.5	4	19.0	1	6.7	11	32.2	0	0.0	11	32.2
AJCC Path Stage	6	54.5	6	28.6	3	20.0	15	37.7	0	0.0	15	37.7

* Total percentages weighted by stratum size to reflect state total.

Note: Data for the individual strata (high, medium, low) are not shown for the software and combined totals.

Table 4. Records with Discrepancies in Cancer and Staging Fields: Colon/Rectum

	Coding Errors								Software Errors		Combined Errors	
	High		Medium		Low		Total *		Total *		Total *	
Field	**No.**	**%**	**No.**	**%**	**No.**	**%**	**No.**	**%**	**No.**	**%**	**No.**	**%**
Site	3	15.8	5	12.8	7	21.2	15	15.5	0	0.0	15	15.5
Histology	1	5.3	3	7.7	2	6.1	6	6.2	0	0.0	6	6.2
Laterality	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Grade	1	5.3	5	12.8	3	9.1	9	8.4	0	0.0	9	8.4
Dxdate	1	5.3	0	0.0	1	3.0	2	3.1	0	0.0	2	3.1
Tumor Size	5	26.3	5	12.8	10	30.3	20	22.2	0	0.0	20	22.2
Nodes Positive	0	0.0	3	7.7	2	6.1	5	3.5	0	0.0	5	3.5
Nodes Examined	0	0.0	3	7.7	2	6.1	5	3.5	0	0.0	5	3.5
Distant Mets	1	5.3	6	15.4	4	12.1	11	9.7	0	0.0	11	9.7
Summary Stage	3	15.8	7	17.9	8	24.2	18	17.7	0	0.0	18	17.7
T Clinical	1	20.0	6	40.0	2	11.1	9	21.2	0	0.0	9	21.2
N Clinical	3	60.0	2	13.3	1	5.6	6	24.7	0	0.0	6	24.7
M Clinical	2	40.0	2	13.3	2	11.1	6	19.2	0	0.0	6	19.2
AJCC Clin Stage	2	40.0	8	47.1	4	22.2	14	32.2	2	3.0	16	35.2
T Pathologic	4	25.0	3	7.9	7	25.9	14	18.5	0	0.0	14	18.5
N Pathologic	3	18.8	0	0.0	4	14.8	7	11.0	0	0.0	7	11.0
M Pathologic	4	25.0	2	5.3	2	7.4	8	15.1	0	0.0	8	15.1
AJCC Path Stage	3	18.8	2	5.3	4	14.8	9	13.0	0	0.0	9	13.0

* Total percentages weighted by stratum size to reflect state total.

Note: Data for the individual strata (high, medium, low) are not shown for the software and combined totals.

Table 5. Records with Discrepancies in Cancer and Staging Fields: Female Breast

	Coding Errors								Software Errors		Combined Errors	
	High		Medium		Low		Total *		Total *		Total *	
Field	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Site	4	21.1	4	12.9	14	35.9	22	20.4	0	0.0	22	20.4
Histology	4	21.1	2	6.5	5	12.8	11	14.9	0	0.0	11	14.9
Laterality	1	5.3	1	3.2	0	0.0	2	3.8	0	0.0	2	3.8
Grade	1	5.3	2	6.5	2	5.1	5	5.7	0	0.0	5	5.7
Dxdate	0	0.0	3	9.7	1	2.6	4	3.7	0	0.0	4	3.7
Tumor Size	1	5.3	9	29.0	13	33.3	23	17.5	0	0.0	23	17.5
Nodes Positive	1	5.3	1	3.2	5	12.8	7	5.7	0	0.0	7	5.7
Nodes Examined	2	10.5	1	3.2	5	12.8	8	8.4	0	0.0	8	8.4
Distant Mets	0	0.0	1	3.2	1	2.6	2	1.5	0	0.0	2	1.5
Summary Stage	1	5.3	1	3.2	0	0.0	2	3.8	0	0.0	2	3.8
T Clinical	0	0.0	6	54.5	13	59.1	19	26.9	0	0.0	19	26.9
N Clinical	0	0.0	4	36.4	8	36.4	12	17.4	0	0.0	12	17.4
M Clinical	0	0.0	2	18.2	3	13.6	5	7.8	0	0.0	5	7.8
AJCC Clin Stage	0	0.0	3	27.3	7	31.8	10	13.9	3	2.6	13	16.5
T Pathologic	3	15.8	4	13.3	12	34.3	19	18.0	0	0.0	19	18.0
N Pathologic	4	21.1	4	13.3	3	8.6	11	17.3	0	0.0	11	17.3
M Pathologic	3	15.8	1	3.3	2	5.7	6	10.5	0	0.0	6	10.5
AJCC Path Stage	2	10.5	3	10.0	10	28.6	15	13.2	2	1.6	17	14.8

* Total percentages weighted by stratum size to reflect state total.

Note: Data for the individual strata (high, medium, low) are not shown for the software and combined totals.

Table 6. Records with Discrepancies in Cancer and Staging Fields: Cervix

	Coding Errors								Software Errors		Combined Errors	
	High		Medium		Low		Total *		Total *		Total *	
Field	**No.**	**%**	**No.**	**%**	**No.**	**%**	**No.**	**%**	**No.**	**%**	**No.**	**%**
Site	8	27.6	1	10.0	1	9.1	10	18.8	0	0.0	10	18.8
Histology	3	10.3	2	20.0	3	27.3	8	16.1	0	0.0	8	16.1
Laterality	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Grade	1	3.4	0	0.0	1	10.0	2	3.1	0	0.0	2	3.1
Dxdate	4	13.8	1	10.0	1	9.1	6	11.8	0	0.0	6	11.8
Tumor Size	4	13.8	4	40.0	6	60.0	14	29.2	0	0.0	14	29.2
Nodes Positive	1	3.4	1	10.0	2	20.0	4	8.0	0	0.0	4	8.0
Nodes Examined	4	13.8	2	20.0	4	40.0	10	19.5	0	0.0	10	19.5
Distant Mets	1	3.4	2	20.0	1	10.0	4	10.1	0	0.0	4	10.1
Summary Stage	4	13.8	2	20.0	0	0.0	6	14.2	0	0.0	6	14.2
T Clinical	6	60.0	2	40.0	4	57.1	12	51.5	0	0.0	12	51.5
N Clinical	5	50.0	2	40.0	3	42.9	10	44.5	0	0.0	10	44.5
M Clinical	4	40.0	2	40.0	1	14.3	7	34.6	0	0.0	7	34.6
AJCC Clin Stage	4	40.0	2	40.0	4	57.1	10	43.5	0	0.0	10	43.5
T Pathologic	7	30.4	5	55.6	3	30.0	15	39.8	0	0.0	15	39.8
N Pathologic	3	13.0	4	44.4	2	20.0	9	25.8	0	0.0	9	25.8
M Pathologic	5	21.7	3	33.3	2	20.0	10	25.9	0	0.0	10	25.9
AJCC Path Stage	6	26.1	3	33.3	2	20.0	11	28.0	0	0.0	11	28.0

* Total percentages weighted by stratum size to reflect state total.

Note: Data for the individual strata (high, medium, low) are not shown for the software and combined totals.

Table 7. Records with Discrepancies in Cancer and Staging Fields: Bladder

	Coding Errors								Software Errors		Combined Errors	
	High		Medium		Low		Total *		Total *		Total *	
Field	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Site	11	57.9	14	37.8	11	40.7	36	48.5	0	0.0	36	48.5
Histology	1	5.3	5	13.5	2	7.4	8	8.4	0	0.0	8	8.4
Laterality	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Grade	3	15.8	8	21.6	0	0.0	11	15.5	0	0.0	11	15.5
Dxdate	0	0.0	1	2.7	0	0.0	1	0.9	0	0.0	1	0.9
Tumor Size	5	26.3	9	24.3	7	25.9	21	25.6	0	0.0	21	25.6
Nodes Positive	2	10.5	1	2.7	3	11.1	6	7.9	0	0.0	6	7.9
Nodes Examined	2	10.5	0	0.0	3	11.1	5	7.0	0	0.0	5	7.0
Distant Mets	1	5.3	1	2.7	2	7.4	4	4.7	0	0.0	4	4.7
Summary Stage	1	5.3	6	16.2	3	11.1	10	9.9	0	0.0	10	9.9
T Clinical	7	46.7	10	29.4	7	28.0	24	35.8	0	0.0	24	35.8
N Clinical	4	26.7	4	11.8	7	28.0	15	20.4	0	0.0	15	20.4
M Clinical	3	20.0	3	8.8	5	20.0	11	15.2	0	0.0	11	15.2
AJCC Clin Stage	5	33.3	9	26.5	5	20.0	19	27.5	0	0.0	19	27.5
T Pathologic	1	25.0	9	32.1	7	50.0	17	26.6	0	0.0	17	26.6
N Pathologic	0	0.0	7	25.0	3	21.4	10	14.6	0	0.0	10	14.6
M Pathologic	1	25.0	7	25.0	6	42.9	14	22.3	0	0.0	14	22.3
AJCC Path Stage	0	0.0	9	32.1	5	35.7	14	19.8	0	0.0	14	19.8

* Total percentages weighted by stratum size to reflect state total.

Note: Data for the individual strata (high, medium, low) are not shown for the software and combined totals.

Table 8. Records with Discrepancies in Cancer and Staging Fields: Lymphoma

	Coding Errors								Software Errors		Combined Errors	
	High		Medium		Low		Total *		Total *		Total *	
Field	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Site	5	26.3	14	37.8	10	28.6	29	30.7	0	0.0	29	30.7
Histology	3	15.8	0	0.0	7	20.0	10	11.0	0	0.0	10	11.0
Laterality	0	0.0	0	0.0	1	2.9	1	0.4	0	0.0	1	0.4
Grade	1	5.6	7	18.9	3	8.8	11	10.7	0	0.0	11	10.7
Dxdate	1	5.3	1	2.7	3	8.6	5	4.9	0	0.0	5	4.9
Tumor Size	0	0.0	8	21.6	8	23.5	16	11.0	0	0.0	16	11.0
Nodes Positive	7	38.9	11	29.7	16	47.1	34	36.5	0	0.0	34	36.5
Nodes Examined	6	33.3	12	32.4	16	47.1	34	34.7	0	0.0	34	34.7
Distant Mets	1	5.6	1	2.7	1	2.9	3	4.1	0	0.0	3	4.1
Summary Stage	5	27.8	8	21.6	9	26.5	22	25.2	0	0.0	22	25.2
T Clinical	13	100.0	16	55.2	20	76.9	49	77.9	0	0.0	49	77.9
N Clinical	13	100.0	16	55.2	20	76.9	49	77.9	0	0.0	49	77.9
M Clinical	13	100.0	19	65.5	20	76.9	52	81.7	6	7.5	58	89.2
AJCC Clin Stage	8	61.5	13	44.8	14	53.8	35	52.8	5	5.5	40	58.4
T Pathologic	10	100.0	20	69.0	22	78.6	52	74.2	0	0.0	52	74.2
N Pathologic	10	100.0	20	69.0	22	78.6	52	74.2	0	0.0	52	74.2
M Pathologic	10	100.0	21	72.4	22	78.6	53	75.5	0	0.0	53	75.5
AJCC Path Stage	6	60.0	11	37.9	19	67.9	36	46.5	0	0.0	36	46.5

* Total percentages weighted by stratum size to reflect state total.

Note: Data for the individual strata (high, medium, low) are not shown for the software and combined totals.

Table 9. Records with Discrepancies in Treatment Fields: All Sites Combined

	Coding Errors								Software Errors		Combined Errors	
	High		Medium		Low		Total *		Total *		Total *	
Field	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Surgery	33	26.6	58	31.4	46	27.7	137	28.3	1	0.4	138	28.7
Surgery Date	21	16.9	29	15.7	31	18.7	81	16.7	19	3.3	100	19.9
Reason no Surgery	25	20.2	21	11.4	29	17.5	75	16.6	0	0.0	75	16.6
Radiation	4	3.2	6	3.2	7	4.2	17	3.4	3	1.2	20	4.6
Radiation Date	4	3.2	7	3.8	9	5.4	20	3.7	2	0.8	22	4.5
Surg/Rad Sequence	12	9.7	12	6.5	18	10.8	42	8.7	0	0.0	42	8.7
Chemotherapy	2	1.6	10	5.4	15	9.0	27	4.0	0	0.0	27	4.0
Chemotherapy Date	6	4.8	10	5.4	13	7.8	29	5.4	1	0.4	30	5.9
Hormone TX	5	4.0	13	7.0	22	13.3	40	6.4	0	0.0	40	6.4
Hormone TX Date	7	5.6	15	8.1	27	16.3	49	8.0	0	0.0	49	8.0
BRM	0	0.0	3	1.6	5	3.0	8	1.0	0	0.0	8	1.0
BRM TX Date	0	0.0	3	1.6	4	2.4	7	0.9	0	0.0	7	0.9
Other TX	1	0.8	0	0.0	5	3.0	6	0.8	0	0.0	6	0.8
Other TX Date	1	0.8	0	0.0	5	3.0	6	0.8	0	0.0	6	0.8
TX Start Date	24	19.4	35	18.9	35	21.1	94	19.4	41	7.7	135	27.1

* Total percentages weighted by stratum size to reflect state total.

Note: Data for the individual strata (high, medium, low) are not shown for the software and combined totals.

Figure 1. Percent Agreement in Coding Demographic and Treatment Fields by Facility Stratum

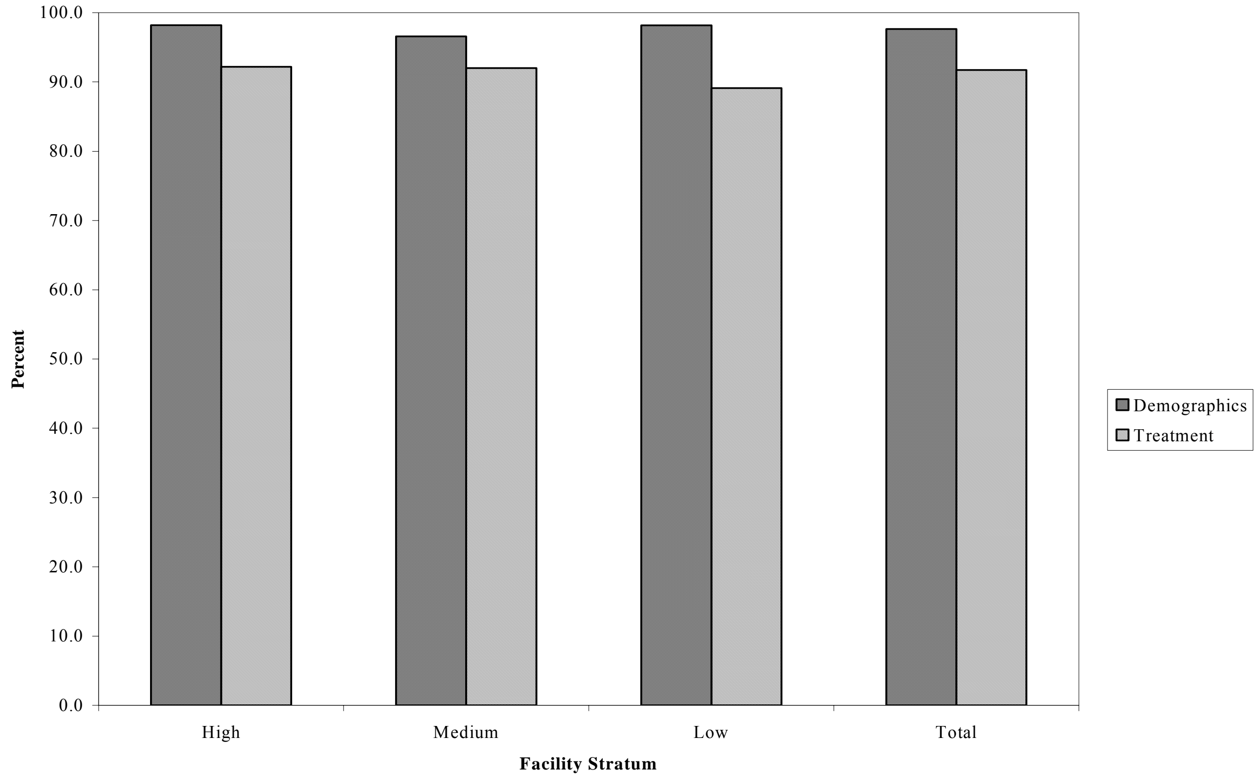


Figure 2. Percent Agreement in Coding Cancer and Staging Fields by Primary Site

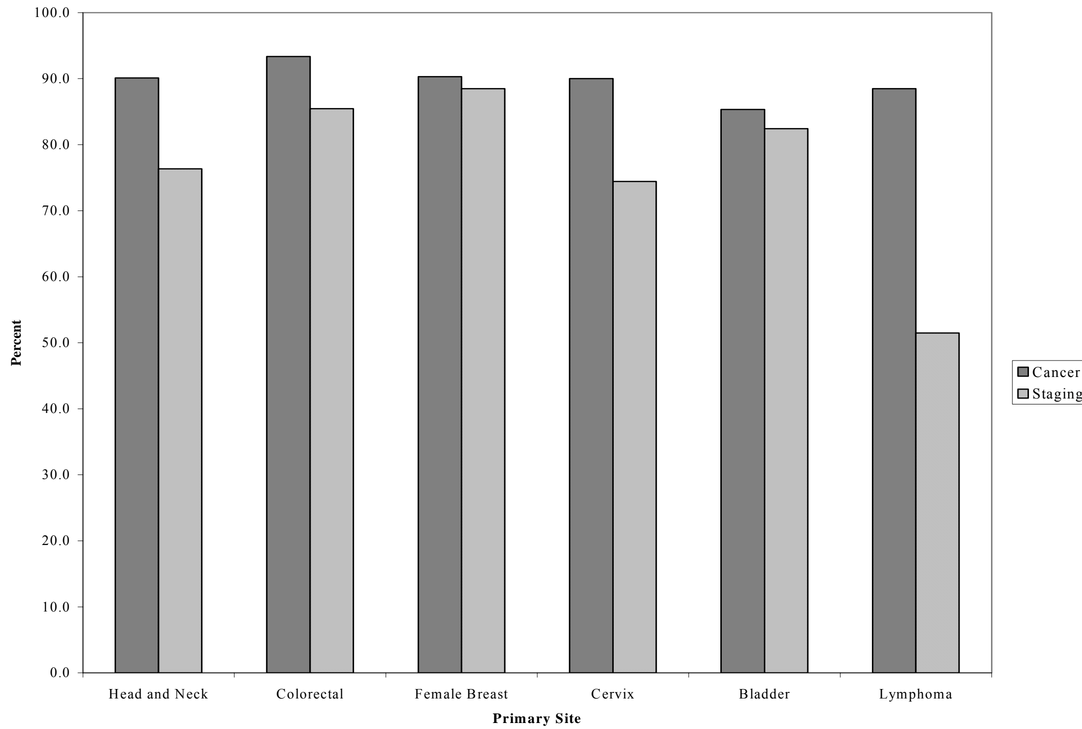


Figure 3. Number of Coding and Software Errors by Demographic and Treatment Fields

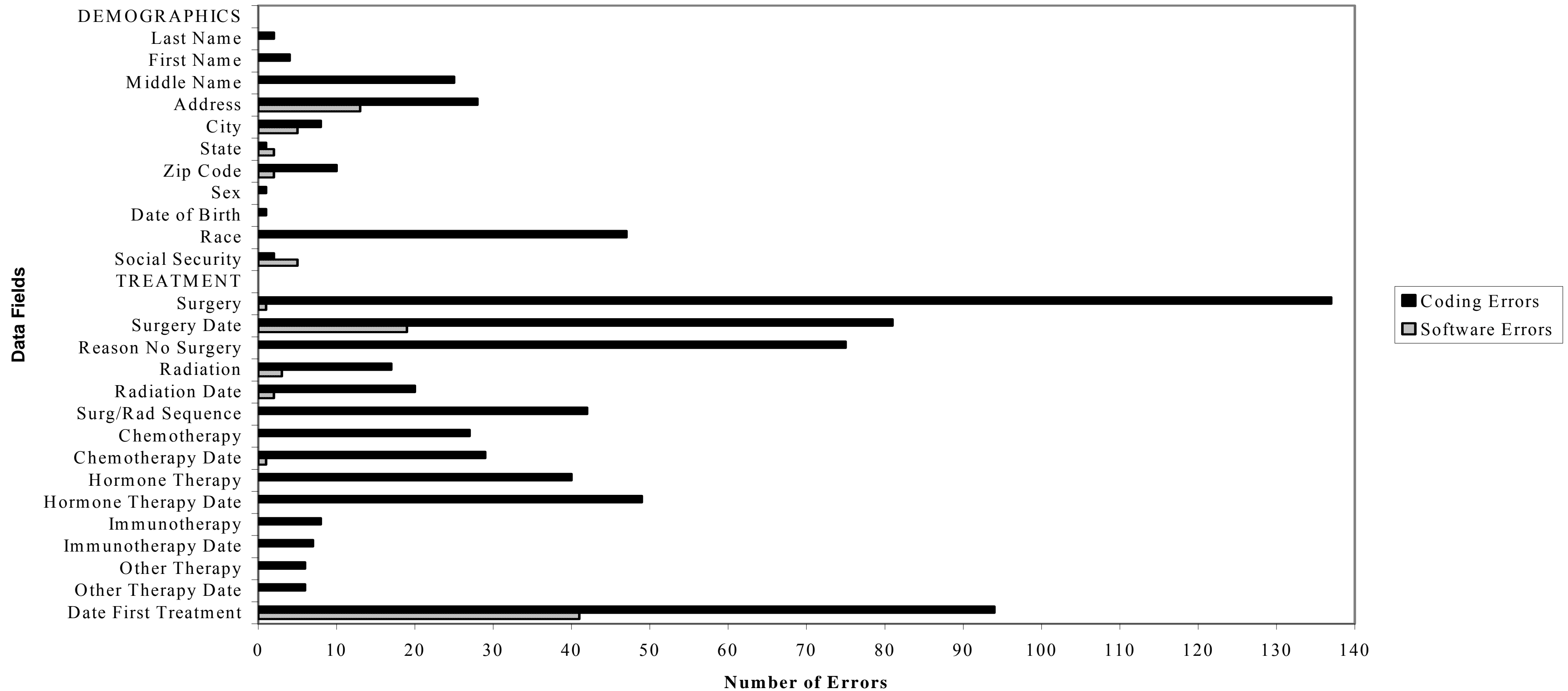


Figure 4. Number of Coding and Software Errors by Cancer and Staging Fields

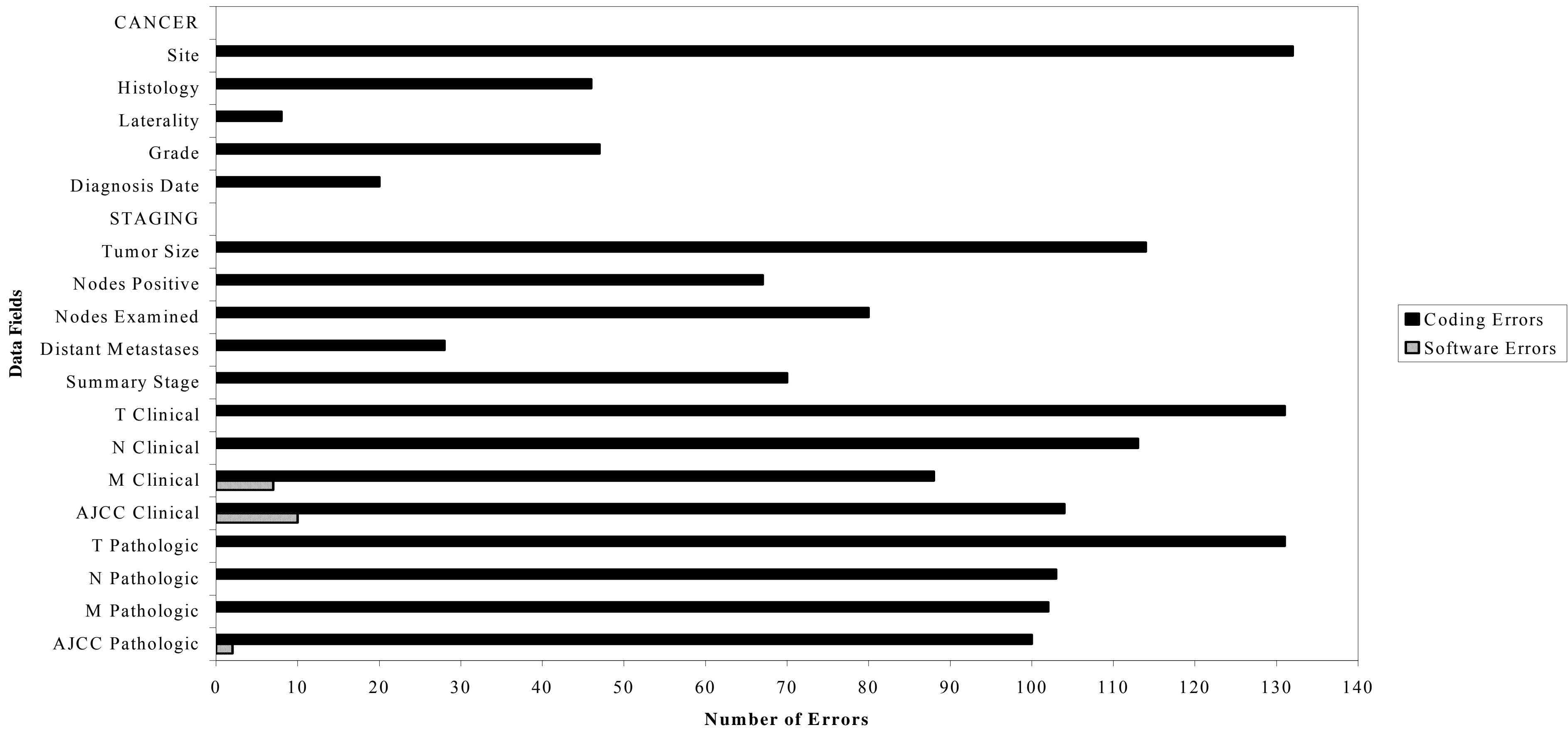
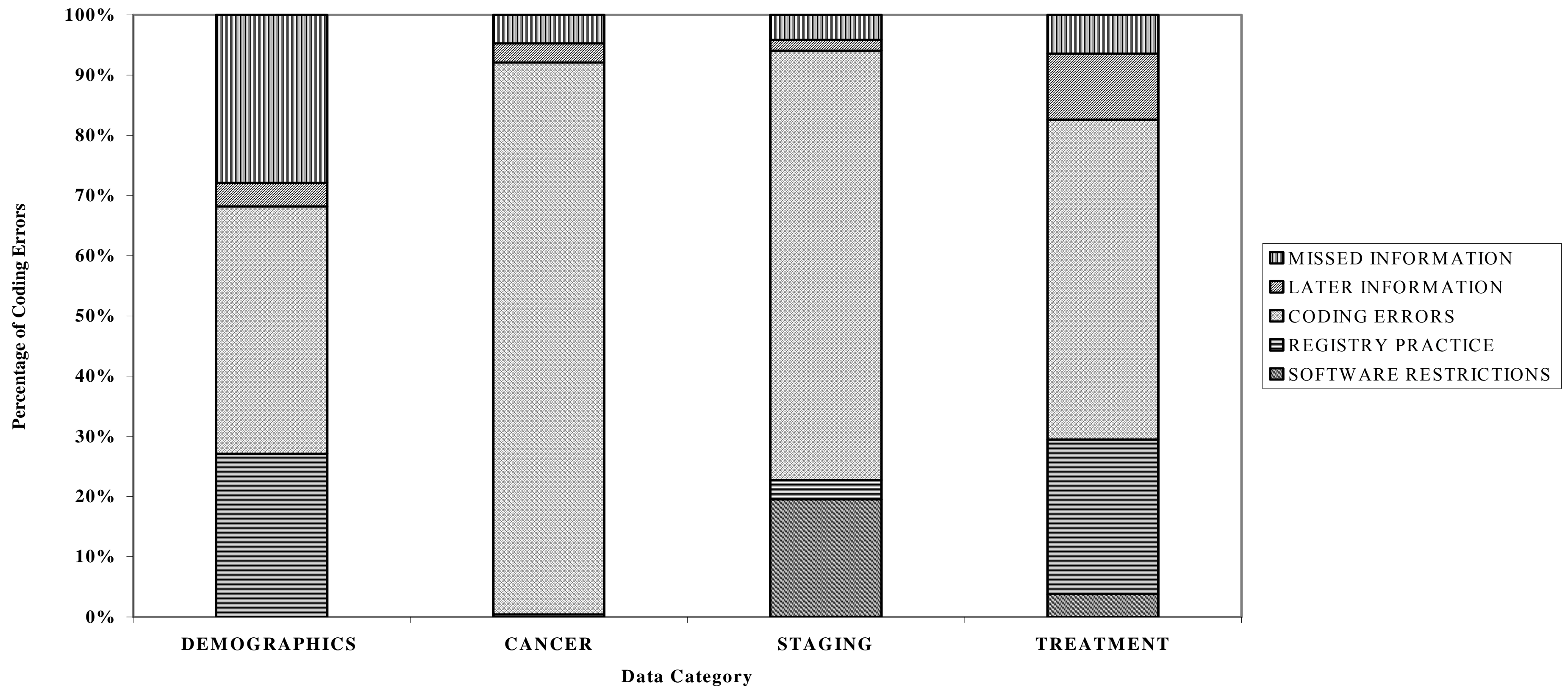


Figure 5. Types of Discrepancies as a Percentage of All Coding Errors by Data Category



REFERENCES

1. Hutchison CL, Roffers SD, and Fritz AG (eds). Cancer Registry Management: Principles and Practice. national Cancer Registrars Association, 1997, p. 155.

2. Menck H and Smart C (eds.). Central Cancer Registries: Design, Management, and Use. Harwood Academic Publishers, 1994, p. 153.

3. Collins EN, Braun JE, and Punyko JA. Evaluation of Abstracting: Cancers Diagnosed in 1997. MCSS Quality Control Report 2000:1.