# Disease Reporting from an Automated Laboratory-Based Reporting System to a State Health Department via Local County Health Departments

Howard D. Backer, MD, MPH[a]
Stanley R. Bissell, MS[b]
Duc J. Vugia, MD, MPH[c]

## SYNOPSIS

**Objective.** The authors assessed the completeness of disease reporting from a managed care organization's automated laboratory-based reporting system to the California Department of Health Services (CDHS) via local public health departments.

**Methods.** The authors identified all positive laboratory tests for 1997 from the computerized database of Kaiser Permanente Northern California for seven infections for which there are statutory reporting requirements: *Campylobacter jejuni*, *Chlamydia trachomatis*, *Cryptosporidium parvum*, hepatitis A, *Neisseria meningitidis*, *Neisseria gonorrhoeae*, and *Salmonella* (*N* = 7,331 reports). Cases were then matched by computer query to records of cases reported to CDHS. To determine why cases were not found in CDHS records, a sample of unmatched cases was searched at two county health departments.

**Results.** Overall, 84.5% (95% CI 83.4, 85.6) of the laboratory reports submitted with accompanying demographic information were successfully matched with cases in the CDHS disease surveillance database. Frequency of matching for specific diseases ranged from 79.4% (95% CI 75.6, 83.3) for *N. gonorrhoeae* to 88.4% (95% CI 85.3, 91.6) for *C. jejuni*. Reports were more likely to be matched when the county of residence was the same as the county of the health care facility. At the county level, reasons for failure of cases to be forwarded to CDHS included: errors due to manual data entry, failure to forward information from the county of diagnosis to the county of residence, and incorrect disease coding.

**Conclusion.** Automated laboratory-based reporting is highly effective, but some data are lost with off-line transfer of information. To optimize surveillance accuracy and completeness, reporting at all levels should be done via direct electronic data transfer.

[a]Immunization Branch, Division of Communicable Disease Control, California Department of Health Services, Berkeley; at the time of this study, Dr. Backer was a Preventive Medicine Resident with the University of California, San Francisco, and Kaiser Foundation Hospitals.
[b]Division of Communicable Disease Control, California Department of Health Services, Sacramento
[c]Disease Investigations and Surveillance Branch, Division of Communicable Disease Control, California Department of Health Services, Berkeley

Address correspondence to Howard D. Backer, MD, MPH, 2151 Berkeley Way, Rm. 712, Berkeley, CA 94704; tel. 510-540-2382; fax 510-883-6015; e-mail <hbacker@dhs.ca.gov>.

Infectious disease surveillance data are used to identify disease outbreaks, track long-term trends, allocate resources, and evaluate the effectiveness of intervention and control programs.[1] The utility of surveillance data depends on the timeliness and thoroughness of reporting, analysis, interpretation, and dissemination of the data.

Passive reporting systems in which health care providers report cases to local health departments are widely used because of their low cost, but underreporting is an acknowledged problem. Reporting rates vary widely by source, with physicians being the least reliable.[2,3] Many factors contribute to low passive reporting rates from physicians, including time restraints, confidentiality concerns, ignorance of reporting requirements, the tendency to report primarily severe diseases, lack of diagnostic laboratory confirmation, the assumption that the laboratory or other office personnel will complete the report, and unfamiliarity with reporting forms.[1–6] Efforts to improve physician reporting rates by providing education, administrative support, and other interventions are modestly successful, if at all, and are usually not sustained.[5,6]

Passive reporting by hospitals is little better than physician reporting.[3,7,8] Standaert summarized reported rates for specific infectious diseases; even for invasive diseases and enteric or respiratory infections widely known to be reportable and considered serious, reported rates in 22 studies ranged from 11% to 93%, with an unweighted median of 42%.[3]

As serology or culture results have become the "gold standard" for many infectious diseases, laboratory reporting has become an increasingly valuable source of surveillance data, with excellent sensitivity and positive predictive value.[9] Laboratory reporting can greatly improve the speed of reporting and the completeness of surveillance data for infectious diseases. Effler found that electronic reporting in Hawaii resulted in a 2.3-fold increase in disease reports to the state and that reports were more complete.[10] Schramm et al.[2] and Vogt[7] evaluated the contribution of laboratory reporting for 11 communicable diseases to the Vermont Department of Health. They found 2,035 reports on 1,636 cases (1.2 reports per case): 71% from laboratories, 10% from nurses, 10% from physician offices, and 9% from other sources. Laboratories reported 100% of the time. For initial reports only, laboratories supplied more than 80% of reports for enteric infections and 48% for hepatitis A and B, but only the minority of cases of invasive *Haemophilus* and meningococcal disease. Harkness compared laboratory diagnosis records to physician and laboratory reporting for shigellosis: 69 of 80 positive cultures (86%) were reported, 3 (6%) by physicians, and 67 (77%) by laboratories.[11] One case was reported by both. Standaert compared passive reporting with an active laboratory-based system over a two-and-a-half-year period for *Neisseria meningitidis* from four counties of Tennessee to the Centers for Disease Control and Prevention (CDC). The passive system identified only about 50% of infections reported by the laboratory.[3]

Passive laboratory reporting may be subject to some of the same limitations as physician reporting. An Australian study found that laboratory reports lagged many weeks behind clinical reports: only half of the laboratory reports were received within three weeks, compared with 98% of physician reports.[12] Problems in the laboratory were related to changes in personnel, equipment, automation, and the assumption that someone else was completing the report. Dominguez et al. found that the wait for reports of laboratory data appeared to slow epidemiological investigation of certain communicable disease cases in Barcelona.[13] Although Vogt found high levels of laboratory reporting in Vermont, the reports often failed to include data on some variables such as patient name, physician name, test date, and result.[7]

Since laboratory reporting is generally more complete than physician reporting, it is now included in the reporting requirements in many states. Title 17, California Code of Regulations, Section 2505 (17CCR 2505) mandates laboratory reporting for 18 diseases. Disease reports are sent to the local or county health department and subsequently forwarded to the California Department of Health Services (CDHS). The California Code of Regulations specifies that laboratories report to the county where the health practitioner works, not the county of residence for the patient.

We evaluated the completeness of laboratory-based disease reporting to CDHS via county health departments by an automated reporting system operated by a regional laboratory of a large health maintenance organization.

## METHODS

We selected seven infections that are reported to county health departments and from there to CDHS under California state regulations: *Campylobacter jejuni*, *Chlamydia trachomatis, Cryptosporidium parvum,* hepatitis A, *Neisseria meningitidis, Neisseria gonorrhoeae,* and *Salmonella.* Physicians are required to report cases of infection with *C. jejuni, N. meningitidis*, and *Salmonella*, while both physicians and laboratories are required to report positive results for *C. trachomatis*, *Cryptosporidium*, hepatitis A, and *N. gonorrhoeae.*

State regulations allow counties to report campylobacteriosis, genital C*hlamydia* infections, gonococcal infections, and salmonellosis to CDHS using summary statistics (simple counts of disease occurrence without demographic data), while counties must file Confidential Morbidity Reports (CMRs) containing detailed demographic information (including personal identifiers) for reported cases of cryptosporidiosis, hepatitis A, or meningococcal meningitis. However, most counties report full CMR data for most of these infections, since they enter the data for their own records.

Kaiser Permanente Northern California (KPNC) is a group-model health maintenance organization that had 2.7 million members in northern and central California in 1997. Results of laboratory tests done either at KPNC medical centers or at a KPNC regional laboratory are recorded in a regional data repository. The computerized system automatically compiles reports of all positive tests for certain infectious diseases, combines them with demographic information from the administrative database, and sends a fax report to the health department in the county where the medical visit occurred. Information reported by fax for each positive test includes patient name, sex, date of birth, address, and phone number; physician name, medical facility, and department; date of specimen collection; source of specimen; test run; and test result. At the time of this study, reports were sent to county health departments weekly; they are now sent daily.

For the present study, analysts from the KPNC Department of Quality and Utilization searched Kaiser's regional database for all positive results for the seven infections for the 1997 reporting year, beginning December 29, 1996. For each positive laboratory test, the following demographics were abstracted from the record: patient name, date of birth, sex; laboratory verification date; and ZIP Code of residence. Files were encrypted and sent electronically via e-mail to the Surveillance and Statistics Section, Division of Communicable Disease Control, CDHS.

At CDHS, one of the authors (SRB) searched state electronic CMR and summary records for the appropriate disease categories and reporting time frame beginning December 29, 1996, and extending to May 30, 1998, to account for any reporting delay. He then imported the files into SAS (Version 7) and purged duplicates. Next, he generated a series of query statements to match the KPNC records with the CMR records at CDHS. First, he found records in which the name, birth date, sex, ZIP Code, and diagnosis matched exactly. He then performed a series of additional queries in which the matching criteria were increasingly relaxed. These subsequent queries allowed variations such as transposition of first and last name, variation between date of diagnosis and verification date, and exclusion of various fields.

When two records were matched by a query, the matched pair was withdrawn from further queries and then verified visually by the investigator. To determine county of residence, we created a table matching ZIP Code with county using ArcView GIS (Version 3.1, 1992–1998, Environmental Systems Research, Inc., Redlands, CA).

We then compared the proportions of KPNC cases that matched with CDHS records by type of infection and by whether the county of residence was the same as the county of the health care facility. These match rates did not take into account cases reported in summary statistics.

Finally, we selected two counties from different parts of the San Francisco Bay Area with high numbers of cases to further investigate cases of four selected infections (*C. jejuni*, hepatitis A, *N. gonorrhoeae*, and *Salmonella*) for which we were unable to match records at CDHS. We sent each of the two county public health departments a list of the unmatched cases with the demographic and laboratory data extracted from the KPNC laboratory database, and the counties searched their records for reports of these cases. Demographic and laboratory data on these unmatched cases were also verified in the KPNC clinical database by one of the authors (HDB).

We used SAS (Version 7) for summary and analytic statistics. Ninety-five percent binomial confidence intervals (CIs) are reported.

## RESULTS

A total of 7,331 laboratory tests positive for the selected seven infections were extracted from the KPNC laboratory database. Of these, 7,213 represented a unique combination of organism and patient demographics.

For cases in which full CMR data were reported from the county ($n$ = 4,090), the frequency of matching with CDHS records was 84.5% (95% CI 83.4, 85.6) (Table 2).

Only 54.3% of these cases matched with CDHS reports using exact spelling of first and last name, birth date, and ZIP Code. The remainder required less restrictive criteria for matching.

Our methodology of matching by demographic and laboratory information did not allow us to match reports sent to CDHS as summary statistics (counts only without detailed demographic or laboratory information; $n$ = 3,123). This was confirmed by a match rate of

**Table 1. Unique cases of selected infections (*N* = 7,213 cases)**

| *Infection* | *Test* | *Site* | *Number of cases* |
|---|---|---|---|
| *Campylobacter jejuni* | Culture | Stool | 1,570 |
| Chlamydia | EIA, FA, Genprobe, culture | Any genital site or urine | 3,624 |
| *Cryptosporidium* | EIA | Stool | 45 |
| Hepatitis A | IgM serology | | 402 |
| *Neisseria gonorrhoeae* | Genprobe or culture | Any site | 1,091 |
| *Neisseria meningitidis* | Culture | Any site | 23 |
| *Salmonella* | Culture | Any site | 458 |

EIA = enzyme-linked immunoassay

FA = fluorescent antibody

only 6.0% for specific infections from counties that usually provide summary statistics for these infections (see shaded areas in Table 3). As a result, only 50.5% of all cases from KPNC matched with cases in the reportable disease database at CDHS.

Match frequency stratified by whether the patient's county of residence was the same or different from the county in which the patient received health care is reported in Table 2**.** Residence and diagnosis in the same county was associated with higher frequency of matching the case at CDHS (relative risk 1.47; $p = 0.0001$), indicating that these cases were significantly more likely to be matched. However, when the relative risk was stratified by infection, the differences were significant only for *Chlamydia*, *C. jejuni*, *N. gonorrhoeae*, and *Salmonella*. No significant difference was found for *Cryptosporidium*, hepatitis A, or *N. meningitidis*.

Table 3 shows the results of the match of laboratory reports to cases reported to CDHS when the county of residence was the same as that of the health care facility**.** Overall, for cases for which the county of residence was the same as that of the health care facility and for which a full CMR report would be expected ($n = 3,383$), the mean frequency of matching was 89.5% (95% C.I. 88.4, 90.5). There was no statistically significant difference in the match rates by type of infection, except for *C. jejuni*, which had a higher probability of a match than other infections based on Cochran-Mantel-Haenszel statistics ($p = 0.001$).

Evaluation of 544 matched reports for which the county of residence was different from the county of diagnosis indicated that 268 (49.3%) were reported to CDHS by the patient's county of residence; 244 (44.9%) were reported by the county where the diagnosis was

**Table 2. Frequency of matching of laboratory-reported cases with cases reported in CMR format to CDHS, by type of infection and by whether county of residence was same as county of health care facility (*n* = 4,090 reports)**

| | *Infection* | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *Patient county of residence* | *Campylobacter jejuni* | *Chlamydia trachomatis* | *Cryptosporidium* | *Hepatitis A* | *Neisseria meningiditis* | *Neisseria gonorrhoeae* | *Salmonella* | *Total* | *95% CI* |
| Same as county in which testing performed | | | | | | | | | |
| Number of reports | 859 | 1,495 | 34 | 342 | 21 | 337 | 295 | 3,383 | |
| Percent matching | 94.8 | 87.8 | 79.4 | 88.9 | 81.0 | 86.4 | 88.5 | 89.5 | 88.4, 90.5 |
| Different from county in which testing performed | | | | | | | | | |
| Number of reports | 161 | 337 | 11 | 55 | 2 | 86 | 55 | 707 | |
| Percent matching | 65.2 | 54.6 | 81.8 | 85.5 | 100.0 | 52.3 | 67.3 | 60.7 | 57.0, 64.3 |
| All reports | | | | | | | | | |
| Number | 1,020 | 1,832 | 45 | 397 | 23 | 423 | 350 | 4,090 | |
| Percent matching | 90.1 | 81.7 | 80.0 | 88.4 | 82.6 | 79.4 | 85.1 | 84.5 | 83.4, 85.6 |

CMR = Confidential Morbidity Report

CDHS = California Department of Health Services

CI = confidence interval

**Table 3. Frequency of matching of laboratory-reported cases with cases reported to CDHS, by county and type of infection for patients diagnosed in county of residence (*n* =5,996 reports)**

| *Northern California counties with KPNC facilities* | *Infection* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | *Campylobacter jejuni* | *Chlamydia* | *Cryptosporidium* | *Hepatitis A* | *Neisseria meningiditis* | *Neisseria gonorrhoeae* | *Salmonella* | *Total CMR matches* |
| Alameda | | | | | | | | |
| Number of reports | 293 | 672 | 3 | 49 | 2 | 256 | 79 | 54 |
| Percent matching | 8.2 | 1.9 | 33.3 | 77.6 | 100.0 | 3.1 | 3.8 | 75.9 |
| Contra Costa | | | | | | | | |
| Number of reports | 173 | 313 | 2 | 30 | 1 | 127 | 45 | 691 |
| Percent matching | 93.6 | 86.3 | 100.0 | 83.3 | 0.0 | 88.2 | 88.9 | 88.4 |
| Fresno | | | | | | | | |
| Number of reports | 32 | 137 | 0 | 4 | 3 | 27 | 9 | 212 |
| Percent matching | 93.8 | 97.1 | — | 100.0 | 66.7 | 88.9 | 77.8 | 94.3 |
| Marin | | | | | | | | |
| Number of reports | 45 | 28 | 4 | 7 | 0 | 10 | 6 | 49 |
| Percent matching | 8.9 | 100.0 | 75.0 | 100.0 | — | 100.0 | 16.7 | 98.0 |
| Napa | | | | | | | | |
| Number of reports | 16 | 19 | 0 | 1 | 0 | 2 | 1 | 39 |
| Percent matching | 100.0 | 100.0 | — | 100.0 | — | 100.0 | 100.0 | 100.0 |
| Placer | | | | | | | | |
| Number of reports | 27 | 23 | 0 | 2 | 0 | 3 | 4 | 59 |
| Percent matching | 85.2 | 82.6 | — | 50.0 | — | 66.7 | 75.0 | 81.4 |
| Sacramento | | | | | | | | |
| Number of reports | 113 | 597 | 3 | 62 | 4 | 216 | 21 | 90 |
| Percent matching | 0 | 0 | 33.3 | 90.3 | 100.0 | 0 | 95.2 | 90.0 |
| San Francisco | | | | | | | | |
| Number of reports | 136 | 185 | 3 | 81 | 0 | 94 | 47 | 267 |
| Percent matching | 96.3 | 0.5 | 100.0 | 88.9 | — | 0 | 66.0 | 88.8 |
| San Joaquin | | | | | | | | |
| Number of reports | 38 | 47 | 0 | 11 | 0 | 10 | 16 | 65 |
| Percent matching | 97.4 | 36.2 | — | 90.9 | — | 20.0 | 87.5 | 93.8 |
| San Mateo | | | | | | | | |
| Number of reports | 124 | 136 | 1 | 21 | 5 | 16 | 38 | 341 |
| Percent matching | 97.6 | 94.1 | 100.0 | 90.5 | 100.0 | 100.0 | 97.4 | 95.9 |
| Santa Clara | | | | | | | | |
| Number of reports | 186 | 546 | 11 | 51 | 3 | 90 | 84 | 971 |
| Percent matching | 95.7 | 83.0 | 100.0 | 98.0 | 100.0 | 75.6 | 96.4 | 86.9 |
| Solano | | | | | | | | |
| Number of reports | 63 | 205 | 4 | 19 | 3 | 53 | 17 | 364 |
| Percent matching | 92.1 | 91.2 | 100.0 | 89.5 | 33.3 | 92.5 | 100.0 | 91.5 |
| Sonoma | | | | | | | | |
| Number of reports | 62 | 68 | 1 | 4 | 0 | 6 | 11 | 152 |
| Percent matching | 93.5 | 85.3 | 100.0 | 100.0 | — | 83.3 | 90.9 | 89.5 |
| Stanislaus | | | | | | | | |
| Number of reports | 2 | 0 | 0 | 0 | 0 | 0 | 2 | 4 |
| Percent matching | 0 | — | — | — | — | — | 0 | 0 |
| Yolo | | | | | | | | |
| Number of reports | 0 | 20 | 2 | 0 | 0 | 3 | 0 | 25 |
| Percent matching | — | 90.0 | 0.0 | — | — | 100.0 | — | 84.0 |
| Total CMR matches | | | | | | | | |
| Number of reports | 859 | 1,495 | 34 | 342 | 21 | 337 | 295 | 3,383 |
| Percent matching | 94.8 | 87.8 | 79.4 | 88.9 | 81.0 | 86.4 | 88.5 | 89.5 |

NOTE: Shaded cells indicate infections reported as summary statistics (without demographic information), for which a match would not be expected. These cases (*n* = 2,613) are not included in the column or row totals.

CMR = Confidential Morbidity Report

CDHS = California Department of Health Services

KPNC = Kaiser Permanente Northern California

made; and 32 (5.9%) were reported by a third county. For the 32 cases that were reported by a third county, 30 (93.8%) had a different ZIP Code on the CMR report from the ZIP Code in the health plan record. By comparison, 80.5% of all matched records had the same ZIP Code on the CMR report as in the health plan record.

Table 4 shows the results of our evaluation of a sample of 95 unmatched cases from two counties. These represent laboratory reports for which there was no record found in the CDHS surveillance database. At least 80% of these records were not found in the case records of the counties either. In five cases, the record was found at the county health department, but apparently had not been forwarded to CDHS. For 14 cases (15%), there was an identified data error that did not allow matching, including four cases in which patients could not be traced back in the managed care organization's database because of name or birth date errors. Seven patients had an incorrect diagnosis due to transcription errors at the county health department or because amended reports were not received by or corrected at the health department.

Evaluation of the four cases of meningococcal disease that did not match revealed that one case was miscoded at the county level as viral meningitis, two cases were miscoded as bacterial meningitis but not specifically as meningococcal, and the fourth record could not be found and may not have been reported to CDHS.

## DISCUSSION

We evaluated the transmission of laboratory reports faxed automatically from a managed care organization to county health departments and subsequently to CDHS. On average, we were able to match 84.5% of positive laboratory reports from the computerized records of KPNC with surveillance data received by CDHS for cases in which counties sent full CMR data. Compared to the findings of other studies, we documented relatively high rates of reporting for these selected infectious diseases, much higher than in most passive systems and comparable to the best active systems.[3,14] Effler found that 90.7% of 357 unique reports to the state of Hawaii were received via electronic reporting, but he estimates that only about 80% of all notifiable laboratory reports could be captured and transmitted electronically by the state's system.[10] Shramm et al. documented the potential of laboratory reporting in the state of Vermont, where reports are sent directly to the state health department,[2] but to our knowledge, no other researchers have evaluated automated laboratory reports that must pass through local health departments to reach the state. Our findings have implications for other jurisdictions in which automated laboratory reports are transmitted through multiple agencies or levels of the public health system.

Assuming that our methodology was valid, and that 100% of cases were reported to a local health department, the confidence intervals in Table 2 imply that

**Table 4. Status of laboratory-reported cases of *Campylobacter jejuni*, hepatitis A, *Neisseria gonorrhoeae*, and *Salmonella* not matched with cases reported to CDHS, two Northern California counties (*n* = 95 cases)**

| *Finding* | *Type of error* | *Number* | *Percent* | *95% CI* |
|---|---|---|---|---|
| Records found in county database | Reporting failure[a] | 5 | 5 | 2, 12 |
| Records not found in county database; residence and medical facility in same county | Reporting and/or recording failure | 46 | 48 | 38,59 |
| Records not found in county of residence database; residence and medical facility in different counties | Forwarding failure[b] | 30 | 32 | 22, 42 |
| Data discrepancy: differences in spelling of name, birth date, diagnosis; laboratory reporting error; different reporting period; or medical record not found in KPNC database | Recording error, data error | 14 | 15 | 8, 2 |

[a]Case apparently not reported to CDHS

[b]Report apparently not forwarded to county of residence

CI = confidence interval

CDHS = California Department of Health Services

KPNC = Kaiser Permanente Northern California

14% to 17% of laboratory reports representing cases of communicable diseases are lost during data transfer between local health departments and the state health department. Although our retrospective analysis did not allow for tracking of these reports with certainty, we were able to identify several points at which data could be lost in the chain of transmission, including failure of local health departments to record cases, failure to forward cases to the county of residence, failure to report cases to CDHS, and recording errors (Table 4).

Almost half of the 95 unmatched cases investigated in two counties were cases not found in the county records, even though the county of residence was the same as the county of diagnosis. This implies that the record, for some reason, did not reach the county health department or was not entered into its computer system. Approximately another third of unmatched cases missing from the county records were cases in which county of residence and county of diagnosis differed. This suggests a failure of the county where the case was diagnosed to forward the information to the patient's county of residence or a failure of the county of residence to record the case and report it to CDHS.

The chain of reporting from local health departments to CDHS indicates how some of these errors may occur. When the laboratory fax report arrives at the county health department, public health workers verify the report. If the patient resides in a different county, the report is supposed to be faxed to the health department of the county of residence. The county where the patient resides then enters the disease and demographic data on a CMR form, which is transmitted to CDHS electronically (by approximately half of the counties in 1997) or sent via regular mail. Many urban areas, including the San Francisco Bay Area, encompass several counties, and patients frequently use health facilities located outside their county of residence. According to our ZIP Code data, 1,223 patients (16.9%) lived in a county different from the one in which they obtained health care.

The location of a health care facility is more reliable and stable than a residential address, making it easier for a laboratory to report, but this creates logistical problems for counties. Although most counties claimed to fax all reports received to the appropriate local health department, an analysis of individual cases suggested otherwise. At least one urban county health department mistakenly assumed that laboratories routinely faxed reports to the county of residence, so the health department did not routinely do so. On the other hand, the relatively high proportion of cases reported from the county of diagnosis when the county of residence was different (244/544 [44.8%]) indicates either inaccuracy in our determination of county by ZIP Code or reporting by local health departments without regard to the county of residence.

Manual data entry may be the most important reason for failure of local health departments to report a case appropriately to CDHS. Data errors may also result from failure to record name or address changes in the clinical record. Our matching technique allowed for differences in more than one demographic data field, but other errors could have prevented us from matching laboratory records to CDHS records. Some diagnoses may be coded incorrectly, as in the case of *N. meningitidis*. Corrections made to erroneous laboratory reports may not be transmitted or may not be linked to the original report.

Limitations of our study design and methodology could also account for our failure to match laboratory-reported cases with state-level surveillance data. Although we attempted to trace back some cases to determine the reason for failure to match, we did not perform a systematic analysis of all unmatched cases to determine errors. Some of the unmatched cases could have been lost due to failure of the automated fax system from KPNC to the counties. Effler found that an electronic data extraction and transmission system failed to function on 30% of the days during the study period.[10] The high match rate found in the present study suggests a much better transmission success than 70%.

We assumed that the laboratory data represented the gold standard, although we found that small proportions of both reporting errors and patient demographic errors originated with the laboratory or health plan. The ZIP Code program that assigned the county of residence to the patients has an unknown margin of error, especially for patients who lived in new communities or near the border of two counties. Some inaccuracy is likely in adjacent counties with rapid growth, especially in a highly urbanized area like the San Francisco Bay area that consists of eight contiguous counties.

While KPNC's automated laboratory data were optimal because they were linked with administrative data to supply address and phone number and other demographics to generate a CMR form, many laboratories currently do not have the necessary demographic information available for a complete report. Other data issues that need to be addressed with regard to electronic disease reporting are duplicate reports, false positive tests, the accuracy of manual data entry, and complete and accurate transmission of data between

counties and from the counties to the state health department.[7,10,15–17]

Laboratory automation and computerization of health care data and electronic linkage have the potential to overcome these data limitations and to further increase the ease, accuracy, and speed of reporting.[1,10,15] Automated laboratory reporting can meet the published criteria for integrating information systems[18] by building on data systems that have already been established, minimizing the collection and reporting burden, and minimizing the effort required to reformat, transmit, and share data with users. Additional anticipated benefits of electronic laboratory reporting include: (*a*) elimination of repeat entry; (*b*) ability to generate timely national summary statistics; and (*c*) increased interaction among public health officials.[19]

Local and national efforts are underway to develop improved electronic disease reporting via a common format that can be transmitted vertically up the levels of the public health system.[16,19–21] Although several successful models already exist, the challenges to implementing a national system of electronic laboratory reporting are substantial.[22] The CDC has sponsored a series of meetings to explore the technical, policy, educational, leadership, and coordination needs to proceed with this effort.[17] The future vision is that public health data will be gathered electronically from various sources including laboratories and health care facilities without time delay and will be used to monitor the health status of populations and guide public health interventions and policies. In the meantime, health departments at the local and state level should maximize the opportunity to exchange data with managed care organizations like KPNC. Sharing of data systems would allow many advantages to both health plans and public health agencies, such as improved vaccination tracking and implementation of preventive health measures.[23]

Our data support the development of nationwide electronic laboratory reporting systems by demonstrating that direct reporting can improve the number of cases reported over the traditional passive system; direct electronic reporting is optimal to avoid the loss of data that results from off-line transfer of information.

## REFERENCES


1. Barrett S, Lau Y. Incompleteness of statutory notification of bacterial gastro-intestinal infection. Public Health 1997;111:183-5.
2. Schramm M, Vogt R, Mamolen M. The surveillance of communicable disease in Vermont: who reports? Public Health Rep 1991;106:95-7.
3. Standaert S, Lefkowitz L, Horan J, Hutcheson R, Shaffnes W. The reporting of communicable diseases: a controlled study of *Neisseria meningitidis* and *Haemophilus influenzae* infections. Clin Infect Dis 1995;20:30-6.
4. Konowitz P, Petrossian G, Rose D. The underreporting of disease and physicians' knowledge of reporting requirements. Public Health Rep 1984;99:31-5.
5. Rothenberg R, Bross D, Vernon T. Reporting of gonorrhea by private physicians: a behavioral study. Am J Public Health 1980;70:983-6.
6. Weiss B, Strassburg M, Fannin S. Improving disease reporting in Los Angeles County: trial and results. Public Health Rep 1988;103:415-21.
7. Vogt R. Laboratory reporting and disease surveillance. J Public Health Manage Pract 1996;2:28-30.
8. Davis SF, Strebel PM, Atkinson WL, Markowitz LE, Sutter RW, Scanlon KS, Friedman S, Hadler SC. Reporting efficiency during a measles outbreak in New York City, 1991. Am J Public Health 1993;83:1011-5.
9. Klaucke D, Buehler J, Thacker S, Parrish R, Trowbridge F, Berkelman R. Guidelines for evaluating surveillance systems. MMWR Morb Mortal Weekly Rep 1988;37:1-15.
10. Effler P, Ching-Lee M, Bogard A, Ieong MC, Nekomoto T, Jernigan D. Statewide system of electronic notifiable disease reporting from clinical laboratories: comparing automated reporting with conventional methods. JAMA 1999;282:1845-50.
11. Harkess J, Gildon B, Archer P, Istre G. Is passive surveillance always insensitive? an evaluation of shigellosis surveillance in Oklahoma. Am J Epidemiol 1988;128:878-81.
12. Thompson S, MacEachern A, Stevenson E. Impact of improved laboratory compliance on notification of genital *Chlamydia trachomatis* infection in Victoria. Sex Transm Dis 1997;24:84-9.
13. Dominguez A, Canela J, Salleras L. Inclusion of laboratory test results in the surveillance of infectious diseases. Int J Epidemiol 1991;20:290-2.
14. Smucker D, Thomas J. Evidence of thorough reporting of sexually transmitted diseases in a southern rural county. Sex Transm Dis 1995;22:149-54.
15. Kilbourne E. Informatics in public health surveillance: current issues and future perspectives. MMWR Morb Mortal Weekly Rep 1992;41:91-9.
16. Parsons DF, Garnerin P, Flahault A, Gotham IJ. Status of electronic reporting of notifiable conditions in the United States and Europe. Telemed J 1996;2:273-84.
17. Centers for Disease Control and Prevention, Electronic Laboratory Reporting Work Group. Electronic Reporting of Laboratory Information for Public Health, January 7–8, 1999: summary of meeting proceedings [cited 2001 May 25]. Available from: URL: http://www.cdc.gov/od/hissb/docs/elr-1999.pdf
18. Morris G, Snider D, Katz M. Integrating public health information and surveillance systems. J Public Health Manage Pract 1996;2:24-7.
19. Bean N, Martin S, Bradford H. PHLIS: an electronic system for reporting public health data from remote sites. Am J Public Health 1992;82:1273-76.

20. Birkhead G, Chorba TL, Root S, Klaucke DN, Gibbs NJ. Timeliness of national reporting of communicable diseases: the experience of the National Electronic Telecommunications System for Surveillance. Am J Public Health 1991;81:1313-15.
21. Koo D, Wtterhall S. History and current status of the National Notifiable Diseases Surveillance System. J Public Health Manage Pract 1996;2:4-10.
22. Pinner R, Jernigan D, Sutliff S. Electronic laboratory-based reporting for public health. Mil Med 2000;165: 20-4.
23. Rutherford G, Backer H. Medicaid managed care and public health data. Public Health Rep 1999;114:225-30.