IN THE MEANTIME.... INTERIM SOLUTIONS TO AUTOMATED DATA CAPTURE
Nancy Cole, BS, CTR
Non-Hospital Reporting Coordinator
Missouri Cancer Registry
THANKS TO
SUE VEST, CTR, MCR PROJECT MANAGER

JEANNETTE JACKSON-THOMPSON, MSPH, PhD
OPERATIONS DIRECTOR

DEB SMITH, CTR, QA COORDINATOR

CHETAN WASEKAR, MD

RASHMI KALGI, DDS
This project was supported in part by a cooperative agreement (#U55/CCU721904) between the Centers for Disease Control and Prevention (CDC) and the Missouri Department of Health and Senior Services (DHSS) and a Surveillance Contract between DHSS and the University of Missouri.
“The need to get paper-based forms into a computer-usable format is greater than ever.” Mark Hayes, www.edocmagazine.com

January/February 2001
83% of all documents are forms
$6 billion spent annually on paper forms
As much as $360 billion spent processing them

Source: AIMM - The Association for Information and Image Management
DATA CAPTURE

- **Data capture** means obtaining data for a computer
  - Collecting documents to be typed or scanned in
  - Turning images into editable/searchable text
  - Not the same as document capture
    - Imaging or photocopying
Automated Data Capture

- **Census Bureau** - In 2000, implemented a data capture & management system.
  - Processed more than 154 million documents in less than 170 days, with a >98% accuracy
Automated Data Capture

- **Automated data capture** means obtaining data directly by an input device without using a keyboard.
  - Bar code readers
  - Scanning pictures
  - Using document readers (OMR, MICR, ICR, OCR)
Types of document readers

- **Optical Character Recognition (OCR):** reads text & software recognizes the letters from their shapes
- **Intelligent Character Recognition (ICR):** Good for capturing handwritten data
- **Magnetic Ink Character Recognition (MICR)/Optical Mark Reader (OMR):** checks or standardized tests
Data Capture Forms

- Data capture forms are designed to have computer input data written on them.
  - Data is standardized
  - Form designer & Adobe
  - PDF
# Missouri Cancer Registry

**Missouri Department of Health and Senior Services**

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Reporting Hospital</td>
<td></td>
</tr>
<tr>
<td>2. Accession Number</td>
<td></td>
</tr>
<tr>
<td>3. Personal Hx 1 Year</td>
<td></td>
</tr>
<tr>
<td>3a. Personal Hx 2 Year</td>
<td></td>
</tr>
<tr>
<td>4. Last Name</td>
<td></td>
</tr>
<tr>
<td>5. First Name</td>
<td></td>
</tr>
<tr>
<td>6. Middle Name</td>
<td></td>
</tr>
<tr>
<td>7. Maiden Name</td>
<td></td>
</tr>
<tr>
<td>8. Alias</td>
<td></td>
</tr>
<tr>
<td>9. Address at diagnosis - Number and Street</td>
<td></td>
</tr>
<tr>
<td>10. Address at diagnosis - supplemental</td>
<td></td>
</tr>
<tr>
<td>11. Dx - City</td>
<td></td>
</tr>
<tr>
<td>12. Dx - State</td>
<td>MO</td>
</tr>
<tr>
<td>13. Dx - Zip code</td>
<td></td>
</tr>
<tr>
<td>14. Dx - County</td>
<td></td>
</tr>
<tr>
<td>15. Medical Record Number</td>
<td></td>
</tr>
<tr>
<td>16. Current Address - Number and Street</td>
<td></td>
</tr>
<tr>
<td>17. Current Address - Supplemental</td>
<td></td>
</tr>
<tr>
<td>18. Current City</td>
<td></td>
</tr>
<tr>
<td>19. C - State</td>
<td></td>
</tr>
<tr>
<td>20. C - Zip code</td>
<td></td>
</tr>
<tr>
<td>21. Name of Spouse</td>
<td></td>
</tr>
<tr>
<td>22. Abstracted by</td>
<td></td>
</tr>
<tr>
<td>23. Social Security Number</td>
<td></td>
</tr>
<tr>
<td>24. Telephone Number</td>
<td></td>
</tr>
<tr>
<td>25. Alcohol History</td>
<td></td>
</tr>
<tr>
<td>26. Tobacco Hx</td>
<td></td>
</tr>
<tr>
<td>27. Tobacco Years</td>
<td></td>
</tr>
<tr>
<td>28. Marital Status</td>
<td></td>
</tr>
<tr>
<td>29. Sex</td>
<td></td>
</tr>
<tr>
<td>30. Race Code: Race 1, Race 2, Race 3, Race 4, Race 5</td>
<td></td>
</tr>
<tr>
<td>31. Span/Hisp Origin</td>
<td></td>
</tr>
<tr>
<td>32. Birth Date - [mm/dd/yyyy]</td>
<td></td>
</tr>
<tr>
<td>33. Birthplace</td>
<td></td>
</tr>
</tbody>
</table>
Database

- Created during form development
- Data - exported to preformatted database (NAACCR 10)
- Stored on hard drive
  - Can be networked
- Exported
  - Data exported to database when verification completed
Other Features

- Can scan supporting documents along with form. Will be bundled as one document.
- Default confidence settings
Phases of Data Capture

- Scanning & Importing
  - (scanner)
- Recognition
  - (reader)
- Verification
  - (verifier)
How it works

1. Incoming Faxed forms
   - Fax Modem
   - OR Form Scanner

2. Form Images
   - Form ID
   - Form recognition

3. Form Data
   - Digitized form data [uncorrected]
   - MS-Access

4. Digitized form data [corrected]
   - MO-CRIS database [Oracle]
   - verifier

5. Edits

   - designer
   - reader
Storage of non-hospital documentation
Document management

“Digital Document Management: There is Far More to It Than Meets the Eye”
Sue-Min Lai, Ph.D., MS, MBA
Journal of Registry Management• 2003 • Volume 30 • Number 4
Kansas Central Registry

- Scan
- Index by Facility #, Date of Diagnosis, Document Type
- Saved in database accessible by staff
Benefits

- Enhances data quality due to reduced data entry errors
- Reduces personnel necessary for data entry
  - 1,800 abstracts from manual hospitals
- Hospitals like the PDF forms
  - Prefer typing to writing
  - Drop-down boxes helpful
- Good scalability – different forms – all being handled by the same process
Issues

- Handwritten text hard to recognize
- Data quality review by non-CTR is clerical only
- Difficulties with lost/missing forms submitted via internal fax
- IT Support
- Unused capabilities
  - Can we use it like KCR?
- Expense
<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Last Contact Date/Death</td>
<td>01/24/2001</td>
</tr>
<tr>
<td>Vital Status</td>
<td>1</td>
</tr>
<tr>
<td>Tumor Status</td>
<td>1</td>
</tr>
<tr>
<td>Cause of Death</td>
<td></td>
</tr>
<tr>
<td>ICD Rev</td>
<td></td>
</tr>
<tr>
<td>Place of Death</td>
<td></td>
</tr>
</tbody>
</table>

### Diagnostic Procedures:
- **PE:**
  - Cough, Sobs, Breath Sounds reduced
- **X-Rays:**
  - CHEST, NSC, CTHEDO
- **Scopes:**
  - Bronchoscopy+
- **Lab:**
  - CEA+2

### Remarks
- Poor prognosis

### Path TNM:
- PT: 4
- PN: 1
- PM: 0

### Clinical-TNM:
- CT: 
- CN: 
- CM: 
- 100. Clin Stage Group: 
- 101. C Staged by: 
- 102. TNM Ed: 

### MA MGHL 80~1 BREMSOUNDS REDUCED
### 95. Diagnostic Procedures:

**PE**
- Cough, Sob, Breath Sounds Reduced

**X-Rays**
- CXR
- CHE, BNS, CTHE, CTHEO

**Scopes**
- Bronchoscopy+

**Lab**
- CEAE

**Remarks**
- Poor prognosis

### 96. Path TNM:
- PT: 4
- PN: 1
- PM: 0

### 97. P Stage group
- 4

### 98. P Staged by

### 99. Clinical-TNM:
- CT: 
- CN: 
- CM: 

### 100. Clin. Stage group

### 101. C Staged by

### 102. TNM Ed

**COUGH, SOB, BREATH SOUNDS REDUCED**
Look before you leap!

- How accurate are paper forms processing systems?
- How easy are electronic forms to design & use?
- How much does it cost to implement these capture solutions & what is the payback?
- How much have the systems improved since you last looked at them?
Can you really use this?

- What are your needs?
- What are your goals?
- What are your considerations for:
  - Turnaround time
  - Accuracy
  - Price
  - Volume
  - Speed
MCR

- **Hardware**
  - Dedicated server
  - Scan station
  - Fax modem
  - Work station

- **Software**
  - Cardiff Teleform Suite (now Verity, Inc.)
  - Phonetools
  - Adobe Professional Version 6.0
Nancy Cole, BS, CTR
Missouri Cancer Registry
PO box 718
Columbia, MO 65205
1-800-392-2829
Colen@health.missouri.edu

*Chetan Wasekar, MD
cjw5f5@mizzou.edu