Biostatistics for Health Care Researchers: A Short Course

Basics of Clinical Data Management

Presented by:

Tim Breen, Ph.D., M.S., C.C.D.M.

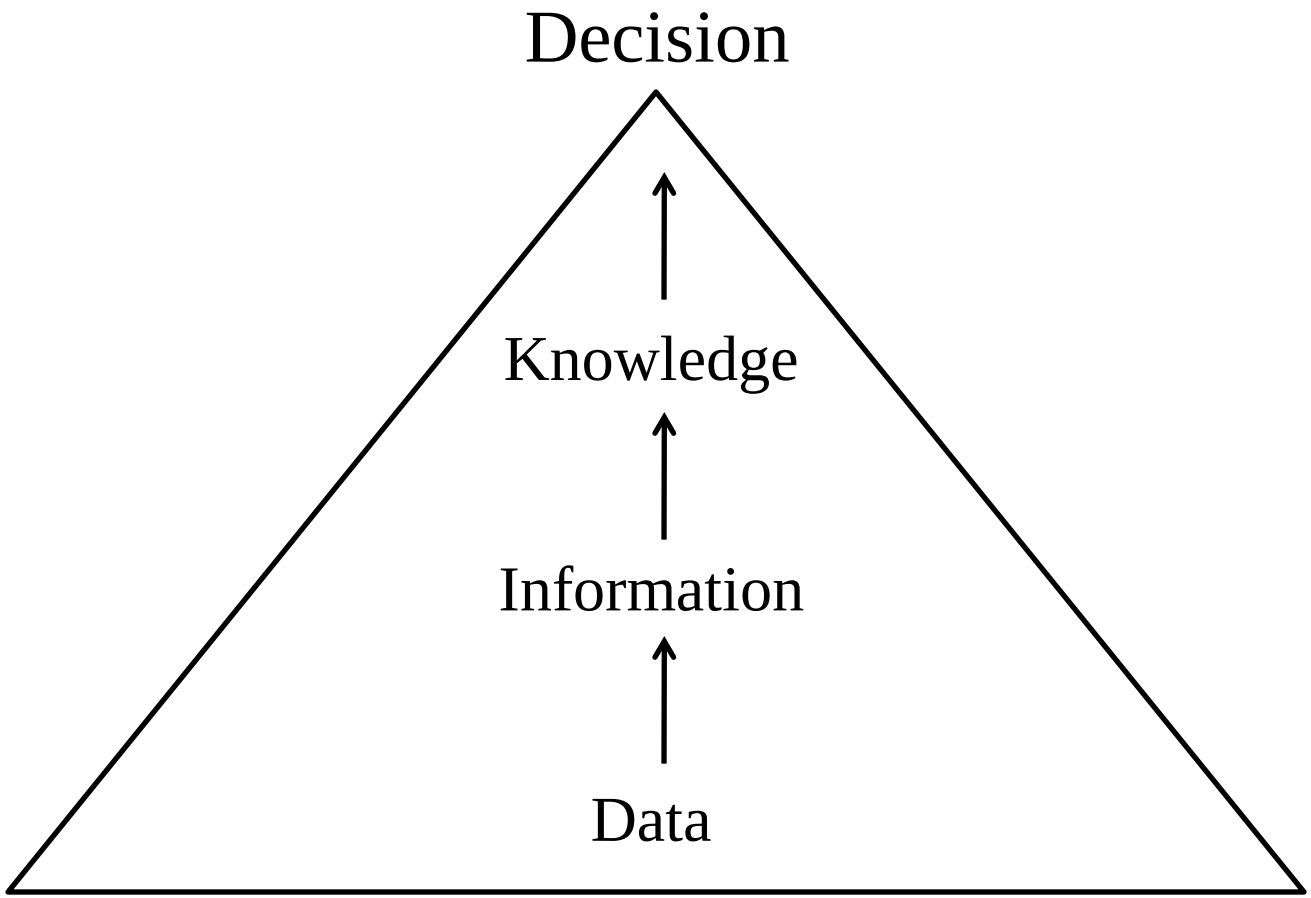
Division of Biostatistics

Indiana University School of Medicine

Objectives

- Introduce the informatics process and data quality
- Describe data management processes
- Describe the role of data management in clinical research

Informatics Process



Data Quality

“High-quality data may be defined as data strong enough to support conclusions and interpretations equivalent to those derived from error free data.”

Assuring Data Quality and Validity in Clinical Trials for Regulatory Decision Making
Workshop Report
Jonathan R. Davis, Vivian P. Nolan, Janet Woodcock and Ronald W. Estabrook, *Editors*
Institute of Medicine
National Academy Press
Washington, D.C. 1999

Research Proposal

- Description of the data management system
- Estimate of the data management budget
- Sample size, proposed data to be collected, data collection schedule

Protocol

"Protocol: All expedited and full review research applications must include a protocol separate from the Summary Safeguard Statement in order to receive IRB review."

IUPUI/CLARIAN

INSTITUTIONAL REVIEW BOARDS

INSTRUCTION PACKET

v05/01/09

For Submitting Applications to the IRB

The Principles of ICH GCP

“2.5 Clinical trials should be scientifically sound, and described in a clear, detailed protocol.”

ICH

International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) – a joint initiative involving both regulators and industry as equal partners in the scientific and technical discussions of the testing procedures which are required to ensure and assess the safety, quality and efficacy of medicines.

ICH

- Founded April 1990
- Tripartite – United States, European Union and Japan
- Six Founding Parties – voting members
- ICH Steering Committee - IFPMA
- Guideline For Good Clinical Practice (GCP) E6
- MedDRA – Medical Dictionary for Regulatory Activities

Stages of a Clinical Trial and DM

- Design and Development

  - Data Management Plan
  - Data Collection Tools/ CRF design
  - Data Management System planning and implementation

- Patient Accrual and Data Collection

  - Ongoing Quality Control
  - Ongoing Trial Monitoring
  - Interim Analysis datasets
  - Reports

- Follow Up and Analysis

  - Ongoing Monitoring/QC
  - Database completeness
  - Database close-up
  - Final Analysis datasets
  - Study Closure

Data Management Plan (DMP)

- DMP – a document which describes and defines all data management activities
- DMP - helps an organization develop and standardize data management procedures

Data Management Plan

Association for Clinical Data Management

- Administrative
- Study Personnel
- Study Objectives and Design
- Timelines and Key Activities
- Database Design
- Monitoring/Validation Guidelines
- Data Flow and Tracking

Data Management Plan

Association for Clinical Data Management

- Data Entry Procedures
- Specification for Clinical Laboratories
- Electronic Data Transfer
- Query Handling
- Backup and Recovery Procedures
- Archiving and Security
- Contract Research Organizations

Data Collection Tools/CRF Design

- Most important step in ensuring data quality is appropriate form design
- CRF Content
- CRF Layout

CRF Content

- CRF questions, prompts and instructions should be clear and concise
- Avoid open-ended questions
- Phrase questions in the positive in order to avoid confusion
- Use appropriate, mutually exclusive responses
- Include units of measurement (e.g. DLCO ml/min/mmHg)

CRF Content

- Collect raw data versus derived data
- Explicitly identify data (e.g. first name, middle name, last name versus name
- Avoid referential and redundant data points
- Include an identifier for the protocol version
- Keep subject identifiers to minimum

CRF Content

DON'T

Q1. How much pain have you experienced lately?

_____

Q2. Weight:_____kg/lbs

Q3. Current Medications:

1. _____
2. _____
3. _____

DO

Q1. How much pain have you experienced lately?

none a little some a lot

Q2. Weight:_ _ _._ kg

Q3. Current Medications?

Tylenol Yes No

Advil Yes No

Aleve Yes No

CRF Layout

- Place key data used in the analysis prominently on the page
- Create well-ordered, structured , easy to follow CRFs
- Adopt consistent style for all the CRFs in the study
- Design the CRFs to follow the data flow from the perspective of the person completing the CRF
- Pilot the CRFs prior to study initiation

CRF Layout

Updated: 03/26/07

HIV-1 RNA RESULTS

IMAGING STUDY OF HIV CEREBRAL INJURY

MRHC 001 / Navia
VI8900

Patient Number ☐☐ ☐☐☐☐ ☐ Date of Specimen collected ☐☐ / ☐☐ / ☐☐☐☐

Form Week ☐☐☐ Completed By (Initials): ☐☐☐

*Sequence Number ☐ * Enter a "1" if this is the first of this form for this date. Designate subsequent forms on the same date with a 2, 3, etc.

Collection Site ☐ UCSD ☐ UCLA ☐ UCLA Harbor ☐ Rochester ☐ Pittsburgh ☐ Penn ☐ Colorado ☐ Stanford

INSTRUCTIONS: Key this form immediately upon receipt of the results from the laboratory.
Use the same header information that was used on the Virology Specimen Tracking Form (VI14688).

1. Was the HIV-1 RNA result obtained from testing lab?..... (1 - Yes, 2 - No) ☐

   If Yes, go to question 2.
   If No, complete (a.) and STOP.

   Reason from testing lab that results were not obtained: [70]

   a. _____

2. Enter the specimen ID number [15]:..... ☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐

CRF Layout, continued

3. Enter the name of the testing lab..... _____

4. Date assay performed:..... ☐☐ / ☐☐ / ☐☐☐☐

5. Type of assay:..... ☐

1 - Roche RT-PCR (Amplicor™) HIV-1 Monitor
2 - Roche Ultra-sensitive ™
3 - Chiron 1 Generation bDNA
4 - Chiron 2 Generation bDNA (ultra-sensitive)
5 - Organon (Teknika™) NASBA
6 - Organon (Teknika™) Nuclisens
7 - Roche RT-PCR (Amplicor™) HIV-1 Monitor V.1.5
8 - Roche Ultra-sensitive V.1.5

9 - Other, specify [30]: _____

6. Enter the plasma HIV-1 RNA result in copies/mL and the quantifier code:

Plasma RNA Results (copies/mL) ☐☐☐☐☐☐☐☐ **Quantifier Code 1 (=), 2 (>), 3 (<)** ☐

7. Enter the CSF HIV-1 RNA result in copies/mL and the quantifier code:

a. Date CSF RNA performed ☐☐ / ☐☐ / ☐☐☐☐

b. CSF RNA Results (copies/mL) ☐☐☐☐☐☐☐☐ **Quantifier Code 1 (=), 2 (>), 3 (<)** ☐

8. Were there any additional comments, (i.e. censor codes) from the testing lab?..... (1 - Yes, 2 - No) ☐

If No, STOP.
If Yes, complete (a.)

a. Comments [70]: _____

Data Management System

- Selection of hardware and Software
- Database Management System (DBMS)
- Data Dictionary
- Database Development
- Data Entry System
- Reporting System
- System Documentation
- System Maintenance and Support
- Security and Data Confidentiality

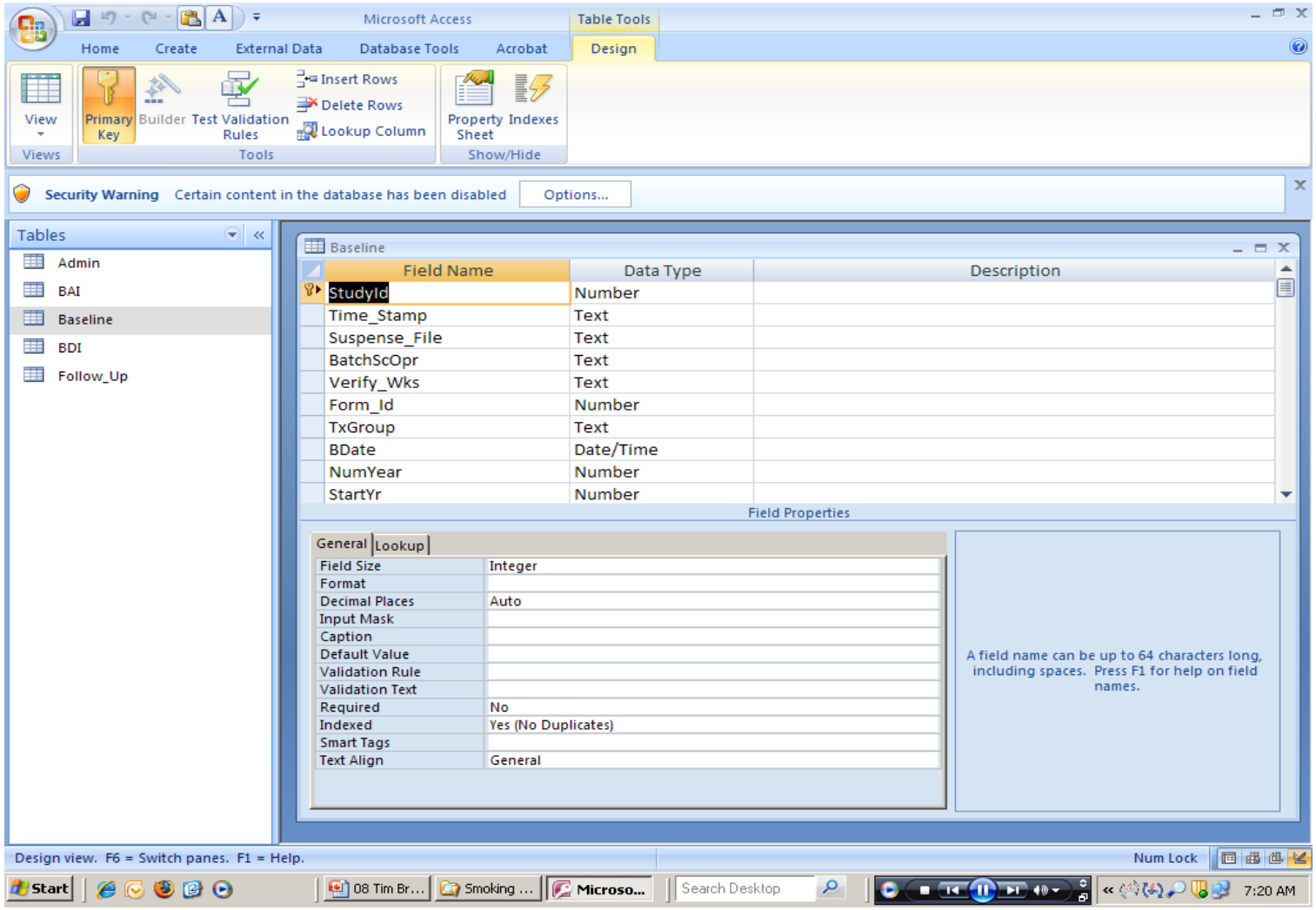
Spreadsheet vs. Database

Property	Spreadsheet	Relational Database
Structure	Cells, Sheets	Tables, Rows, Columns Queries, Reports
Usage	Short Term	Long Term
Data Integrity	Possible, not common	Enforced
Multiple Copies	Easily Duplicated	More difficult
Flexibility	Fewer Uses	Multiple Uses
Concurrency	One user	Multiple users

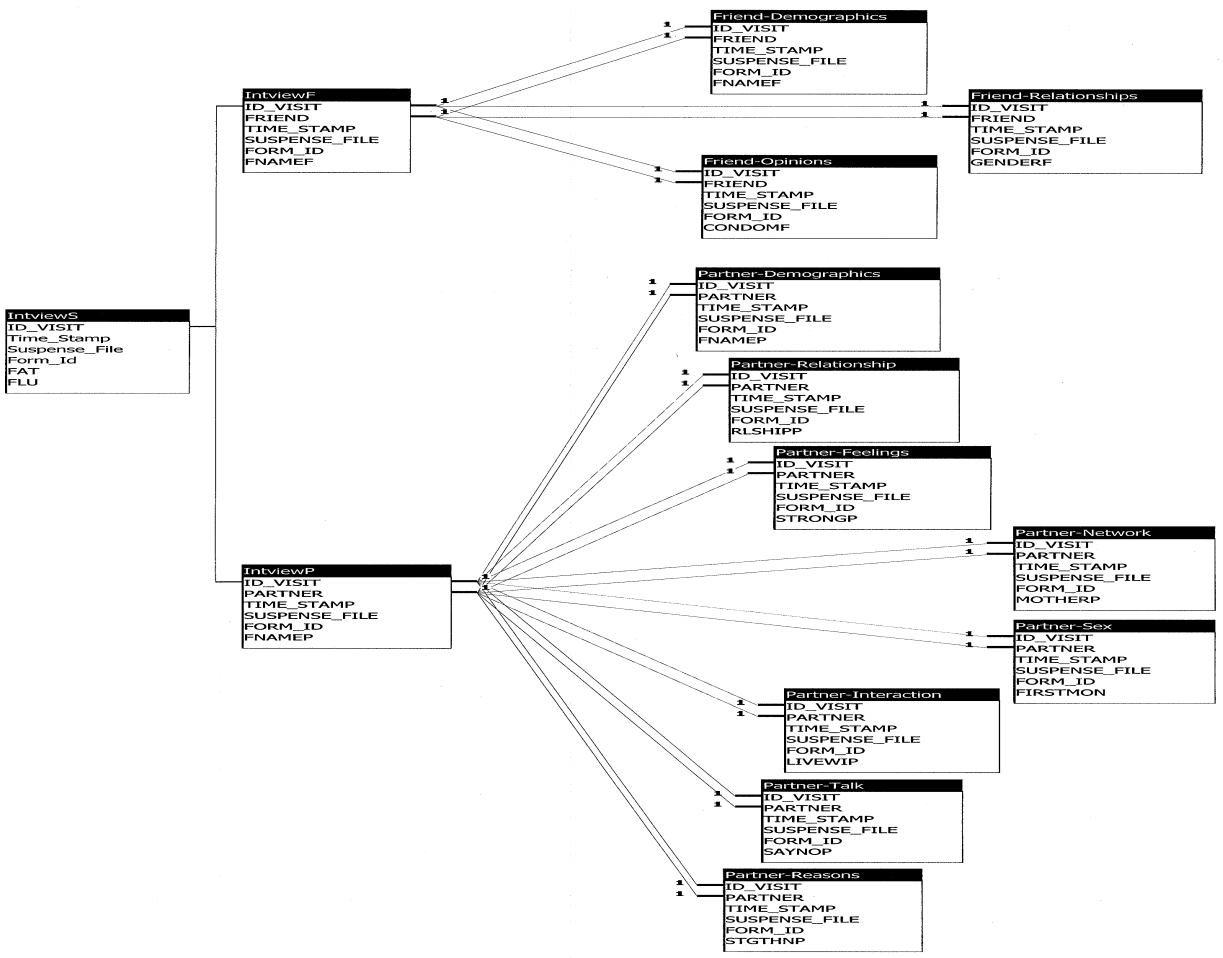
Data Integrity

40128
18212

Data Management System



Data Management System



Database Design

- Identify key fields on the CRF
- Select appropriate data types
- Choose meaningful field names
- Maintain consistency of names and data types for key fields
- Prepare for missing data
- Coding for missing data
- Database tables
- Data dictionary

Database Validation

- Test data entry screens to ensure data are mapped to the correct fields
- Validate the data field definitions in terms of length and type
- Verify that out-of-range data are flagged and error messages trigger properly
- Verify that primary key fields are assigned correctly, no duplicates
- Validate edit, range and logic checks

Database Validation
continued

- Verify algorithms and all other study specific programming
- Record date, time and user stamps for audit trail purposes
- Verify the proper function of data transfers, uploads, exports or integration of external data sources

Quality Assurance

All those planned and systematic actions that are established to ensure that the trial is performed and the data are generated, documented (recorded), and reported in compliance with Good Clinical Practice (GCP) and the applicable regulatory requirement(s).

ICH GCP E6 Glossary

Quality Assurance

- QA focuses on error prevention
- QA begins with protocol and CRF design
- QA includes data processing and analysis
- QA examples – quality audit, defining process, selection of tools and training

Quality Control

The operational techniques and activities undertaken within the quality assurance system to verify that the requirements for quality of the trial-related activities have been fulfilled.

ICH GCP E6 Glossary

Quality Control

- QC focuses on process monitoring
- QC should be applied at each stage of data handling to ensure data are reliable and processed correctly
- QC constant process with feedback
- QC examples – double data entry, regular evaluation of error rates, monitoring

Monitoring

The act of overseeing the progress of a clinical trial, and of ensuring that it is conducted, recorded, and reported in accordance with the protocol, Standard Operating Procedures (SOPs), Good Clinical Practice (GCP), and the applicable regulatory requirement(s).

ICH GCP E6 Glossary

Audit

A systematic and independent examination of trial related activities and documents to determine whether the evaluated trial related activities were conducted, and the data were recorded, analyzed and accurately reported according to the protocol, sponsor's standard operating procedures (SOPs), Good Clinical Practice (GCP), and the applicable regulatory requirement(s).

ICH GCP E6 Glossary

Data Privacy/Security

- Strip source documents of personal identifiers
- Genetic data requires special consideration
- HIPAA compliant applications
- Limited access
- Audit trails
- Antivirus software
- Backup and recovery

Data Analysis and Reporting

- IRB Continuing review
- Data Safety Monitoring Boards (DSMB)
- Interim analyses for safety or efficacy
- Interim analyses for abstracts
- Final analyses

Analysis

- Delineate the research question with the statistician and investigator
- Determine critical data required to evaluate the research question
- Prepare statistical analysis datasets and QC listings
- Work with statisticians to complete the analysis
- Participate in the review of the analysis with the statistician and investigator

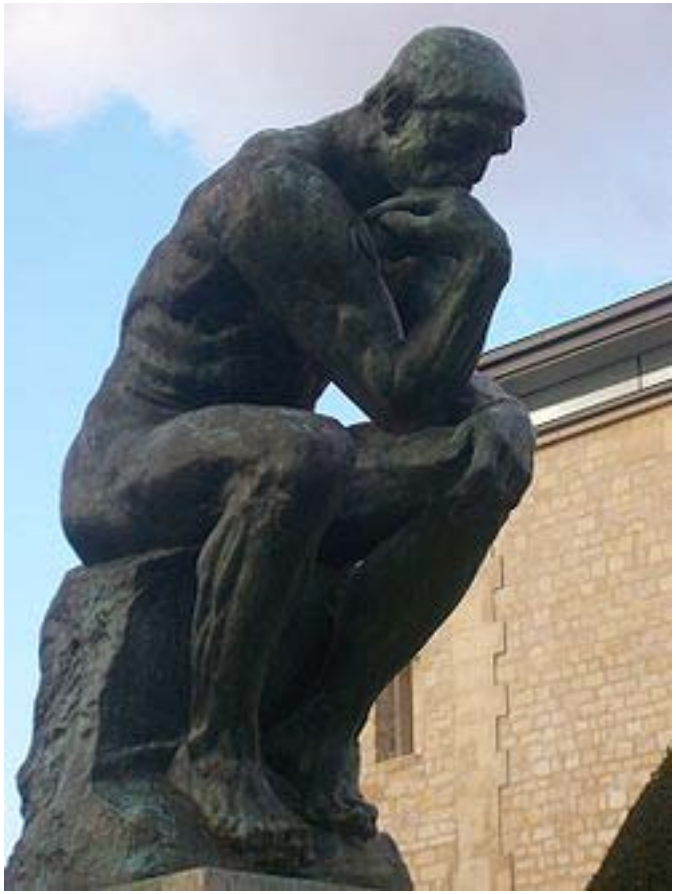
Conclusions

- Data quality is the most important aspect of clinical data management
- Data quality must support the evaluation of study objectives
- Data quality is a multidisciplinary effort
- Data quality requires sufficient resources and expertise

Far and away the best prize life offers is the chance to work hard at work worth doing.

Theodore Roosevelt

Questions



tbreen@iupui.edu

Miscellaneous Verbatims: Coding Challenges

- Went to hell
- Recurrent fatal stroke
- Hears New Age music when the furnace turns on
- LK RTCTL UNSP XTRNDL
- Charcoal-like, gritty granules in his underwear
- Can’t control patient during menses
- His nodule is sticking out
- Normally normal after drinking coffee
- Died of cancer of the placebo
- Superior members fornication
- Barely visible posterior
- Seeing people in room, seeing chickens at window
- Seeing stars and chicken farting
- Patient recently began new job where he works around chicken wings and barbecue sauce