Enriching and Mining Missing Data: Design and Analysis Issues

Greg Soon, Ph.D.
Lead Mathematical Statistician
Division of Biometrics IV
Office of Biostatistics/OTS/CDER/FDA

The Fourteenth Annual Biopharmaceutical Applied Statistics Symposium (BASS)
Nov 5-9, 2007, Savannah, Georgia
Wednesday, November 7, 8-8:50am, Mulberry Inn

Disclaimer

Views expressed here are of the presenter and not necessarily of the FDA

Background (Personal Bias)

- Statistics is not just about the analysis, it is about the process from design, data generation, verification, analysis, to interpretation. All for maximizing the utility of, and drawing the proper conclusions from, the experiment.
- Missing data may not be the real problem, it is the missing information that may be the real problem. Missing could be and often is part of the valid outcome.
- Missing data issue is most efficiently addressed through design.
- A simple and somewhat incorrect method is preferred over a complicated and somewhat incorrect method

Outline

- Missing Data Classification
- Missing, or Not Missing - Fundamental Issues
- Toward More Informative Missing
- Analysis Issues
- Regulatory Role in Missing Data

Definitions

- Webster: Missing = Absent;
  Absent: not present or attending; not existing; lost in thought
- Other dictionaries: Lacking, not found

- Interpretation in clinical trials
  - Subject did not comeback for the clinical assessment, or refused to participate.
  - Subject died
  - Subject absent minded?
  - Subject do not have the target for measurement
  - Lost to follow-up

Missing Data Classification

Missing Data Classification

- In Study Transient Missing
  - Subject remained in the study but did not come to some clinical or lab visits, or failed to fill the diary completely, or some records were deemed not usable
- Lost to Follow-up
  - Subject missed scheduled assessments and did not return for final assessment, the subject could not be contacted.
- Discontinuations and Treatment Changes
  - Subject discontinued or modified the assigned treatment, typically with the knowledge of the investigators. Usually the reasons are documented.
- Deaths

Causes for Transient Missing

1. Holiday visits to relatives, School re-union, Professional meetings, Win lottery, Jury duty, Hurricane, Marriage, Funeral, Car accident, Traffic jam, Too much work waiting, …
2. Lab or technician have problems; Machine malfunction; Undeterminable outcome; Reading errors;
3. Privacy Protection or confidentiality
4. Uncertainty in data due to un-readable handwriting, mistakes in recording, lost record, etc.
5. Due to subject “do not know”, for example, the subject may not be able to recall treatment history
6. Adverse events, tolerability issues, lack of efficacy, feeling well.

Causes for Transient Missing (Cont.)

- Should be rare among hospitalized, nursing home or other closed facilities
- The reasons in cases of 1-5 are often not specified
- May or may not be related to the treatment. In general 1-5 are less likely to be *Directly* related to treatment, but may be related *indirectly*
  - For example, subject may visited a relative during holiday because feeling depressed and need support. Otherwise the subject may have invited the relative home and will not miss the clinical visit.
  - Patient involved in a car accident and missed the visit. The patient was feeling dizzy that day

Causes for Missing Due to Lab Procedure

- Risk in Lab Procedure
  - Fear of blood, fear of pain, fear of the risk in medical or lab procedures like biopsy
  - May occur among hospitalized subjects
  - May or may not be treatment related
  - Only affect selected measures
  - Example: Liver biopsy is invasive and have risk, patients with hepatitis may refuse if they do not feel it is beneficial: they feel they have been doing well so they do not expect to see any worsening in their condition to warrant a change in therapy, or they feel so sick that they know the drug is not helping them.

Other Missing Are Likely Directly Treatment Related

- Lost to follow-ups, permanent discontinuations and deaths could be due to similar reasons,
- But it tend to be more directly treatment related
  - Feel too weak to go, depressed, sleepy, diarrhea, headache, dizzy, or other adverse events
  - Injection or inhalation too difficult, pills taste not tolerable, lab procedure is too difficult, or other tolerability issues
  - Did not achieve meaningful change in lab measures, did not feel any better, did not think the risk of the infection exist, or other lack of efficacy problems
  - Feel too well, feel cured, feel certain not infected (in a prevention trial)

Example: HIV Trial

- Grouping of reasons for discontinuation and lost follow-up in HIV trials based on over 10,000 subjects.
  - Virologic failure
  - No virologic response
  - Lack of efficacy
  - Protocol defined immunological failure
  - Disease progression
  - Worsening of disease under study
  - Worsening of other pre-existing disease
  - Death
  - Adverse events
  - Pregnancy

Example: HIV Trial (Cont.)

- Protocol violation
- Non compliance
- Lost to follow up
- Consent withdrawn
- Refuse treatment
- Fail to return
- Never treated
- Physician’s decision
- Personal reason
- Admin/Other
- Other
- Non-categorized reason
- Reason missing.

Missingness by Mechanism

Let D = {X , Y} be the data matrix, where D includes both independent (X) and dependent variables (Y).

We assume that some elements of the data matrix are missing.

Let M denote the *missingness indicator matrix* with the same dimensions of D. Each element of M is a one or zero that indicates whether or not an element of D is missing.

M_{ij} = 0 indicates that the *i*-th observation for the *j*-th variable is missing, but that the data could be observed.

M_{ij} = 1 means that piece of data is present.

Finally, let D_{obs} and D_{mis} denote the observed and missing parts of the D.

$D = \{D_{obs}, D_{mis}\}$.

MCAR: Missing Completely at Random

If the data are missing completely at random then missing values cannot be predicted any better with the information in D, observed or not.

Formally, M is independent of D. So, P(M | D) = P(M).

A process is missing completely at random if, say, an individual decides whether or not come back for a clinical visit or lab evaluation on the basis of coin flips.

If subjects are more likely to miss clinical visits when they feel well, then the data are not missing completely at random.

In the unlikely event that the process is missing completely at random, then inferences based on listwise deletion are unbiased, but inefficient because we have lost some cases. Other, more efficient methods can be constructed.

MAR: Missing at Random

If the data are missing at random then the probability that a cell is missing may depend on D_{obs}, but after controlling for D_{obs} that probability must be independent of D_{mis}.

In other words, the process that determines whether or not a cell is missing should not depend on the values in the missing cell.

Formally, M is independent of D_{mis}: $P(M | D) = P(M | D_{obs})$

For example, if patients who are doing well on a lab marker (ALT) tend not to have biopsies, and the actual biopsy value has no impact on the decision of not having biopsies after controlling for the ALT. ALT not missing. Then the missing of biopsy is MAR when ALT and biopsy data are grouped together.

If data is missing at random, then inferences based on listwise deletion will be biased and inefficient.

- Multiple Imputation approach will work
- Other modeling approaches may work as well

Non-ignorable Missing

If the probability that a cell is missing depends on the unobserved value of the missing response, then the process is *non-ignorable*.

Formally, P(M | D) cannot be simplified.

Very common is clinical trials.

In treatment trials, patients who are not responding well, going through serious adverse events, or doing extremely well may feel continued treatment or lab visits beneficial.

If your missing data is non-ignorable, then inferences based on listwise deletion will be biased and inefficient (and multiple imputation algorithms won't be of much aid).

Important Observations: Selection Dependant

Let $D^*=\{D, Z\}$,

then $M^*=\{M, M_Z\}$, $D^*_{obs}=\{D_{obs}, Z_{obs}\}$ *and* $D^*_{mis}=\{D_{mis}, Z_{mis}\}$

1. D* is MCAR then D is MCAR
2. MAR will depend on what variables being included
3. We can only model what being recorded, but there may be other unrecorded, or unknown variables that need to be included to achieve MAR – unknown unknowns

Donald Rumsfeld on Missing Data

As we know,
There are known knowns.
There are things we know we know.
We also know
There are known unknowns.
That is to say
We know there are some things
We do not know.
But there are also unknown unknowns,
The ones we don't know
We don't know.

Feb. 12, 2002, Department of Defense news briefing

Missing, or Not Missing – Fundamental Issues

Example

Consider HIV Trials. Assume the trial is designed for 48 weeks, a subject discontinued at Week 24 due to adverse events, and the primary endpoint is suppression of viral load below 50 Copies/mL.

The subject likely will switch to a new treatment

It there a missing data problem?

1. Yes – subject do not have Week 48 on treatment viral load measurement so we can not determine the primary endpoint

2. No – Discontinuation due to AE reflected what will happen in the real world, insisting on having on treatment viral load measurement at Week 48 is like insisting on measuring blood pressure of a dead person – it is meaningless

No Missing Data – Just Need Proper Interpretation

1. Discontinuation due to AE is an outcome, reflecting the consequence of the treatment.
2. The issue is not to find the on treatment viral load of the subject, rather, the issue is how to weigh the two different clinical events: discontinuation vs. suppression of viral load
3. Missing mechanism along is insufficient for such a job, as illustrated by the following example.

HIV Example Continued

1. If the scenario described occurred in a highly treatment experienced population, because the limited treatment options left the consequence could be severe – the patients may soon burn out all the options and lose control of the HIV infection that could lead to disease progression and death
2. In naïve population, the consequence could be much less severe – there are more options left and the patient will have time waiting for new effective interventions

HIV Example Continued

1. The reasons for discontinuation were the same in the two populations, the underlying mechanism could also be the same – for example the events could be MAR when conditioning on AEs

2. But the interpretations could be vastly different
   1. In experienced population discontinuation due to AE is clearly on par with failure of viral suppression
   2. In naïve population discontinuation due to AE may be between success and failure

The Purpose of Imputation

Which question do we want to address?

- Had the subjects come back for visit, what would be their outcome?
- Had the subjects *continued treatment* and come back for visits, what would be their outcome?
- What is the consequences of the treatment strategy to the subjects *in the long run*?

The Purpose of Imputation (HIV Example)

1. Had the subjects come back for visit, what would be their outcome?
    1. The subject may be a success at the end of the trial, but that success is likely due to the new therapy the subject is taken, not due to the originally randomized therapy
    2. This approach will favor the treatment arm who may have more such discontinuations
    3. Could be a reasonable question when no new options exist for these subjects
    4. Could be the right question for mortality or irreversible morbidity endpoints

The Purpose of Imputation

2. Had the subjects *continued treatment* and come back for visits, what would be their outcome?
    1. This is the wrong question to ask. We can not ask a subject to continue a treatment that is not beneficial, and it will not reflect the medical practice after drug approval
    2. Similar to ask what is the blood pressure of a dead person had that person still alive.

The Purpose of Imputation

3. What is the consequences of the treatment strategy to the subjects *in the long run*?
    1. This is the right question, especially when the endpoints are biomarkers or symptoms
    2. In HIV case, such subjects are considered as treatment failures due to the following reasons
        1. Not able to take the drug means there is no future benefits. In fact if no new drugs are introduced to the regimen, discontinuation of therapy will result in quick return of viral load to baseline
        2. Adverse events, especially serious adverse events, are harmful
        3. Previous drug exposure could have introduced resistance virus and reduce the usefulness of future drugs

Reducing Missing Data and Increase Information Contents

It Is Possible to Reduce Missing Data: Examples

1. In a large one year genital herpes suppression trial, the missing rate was 40%. FDA rejected the NDA citing the missing data made the trial not interpretable. Subsequently the trial was repeated and the missing rate was 20%.

2. When the first anti-viral agent, Epivir, was submitted for approval for the treatment of hepatitis B, the studies had missing rates ranging from 15 to 30% for the primary endpoint (liver biopsies). Subsequently FDA sent comments to the sponsors who were to conduct hepatitis B trials, warning that excessive missing will likely make the trials not interpretable. So far all new trials had missing rates 7-15%.

Reducing Missing by Better Planning

Extra efforts by investigators and collaboration from subjects are the key.

- Understanding by all parties that a large trial with excessive missing is worse than a small but clean trial
- Setting up expectation and taking steps to achieve it
- Well planned protocol and investigator brochure having details on what to do under different scenarios
- Better training of the investigators
- Incentives for the investigators and patients for clinical visits
- Use of modern technology

Reducing Missing by Better Execution

Active instead of passive contact with subjects

- Keep a variety of contact information from subjects: telephone, email, family member/guardian, …
- In case a subject failed to return for clinical or lab visit, investigators should contact subjects and encourage them for clinical visit
- If the subject could not come for the scheduled visits, alternative visit may help
- Need to have a clear understanding of the reason for not coming back and the basis for the reasons.
- Information on the general well-being of the subjects will also help

Reducing Missing by Better Off-treatment Follow-Up

End of treatment does not mean end of information

- Information in the off-treatment follow-up could help the interpretation of the data during the follow-up
- Can be used to perform true intent to treat analysis. This is especially useful for mortality or irreversible morbidity endpoints
- Can be done efficiently by following every subject until the last subject complete the study and the minimal required follow-up. This way the trial duration will not be increased and submission time not affected

Reducing Missing by Better Prioritization

Knowing what to collect and what to give up

- Excessive burden on investigators and subjects may be counter-productive
- Prioritize the variables needed. The variables seek should be the ones thought most relevant to the interpretation of the results and achievable
- When large number of missing is expected, a pre-selected subset of subjects should be followed more thoroughly instead of all subjects to make it feasible. This strategy can be refined to make it more informative

Reducing Missing by Better Selection of Endpoint

1. Time to event type endpoint sometimes can be determined based only on early information
2. Coarser endpoint like success/failure could be more powerful than finer endpoint like change from baseline when imputation is considered
3. Coarser endpoint like success/failure could be easier in having credible imputations than finer endpoint like change from baseline

When Will Responder Analysis Be More Powerful Than Change From Baseline?

Minimum Responder Rate of the test arm Required

True Effect	Rate of Discontinuation		
	0%	20%	50%
0.5	61%	16%	8%
1	100%	24%	13%
1.5	100%	39%	20%

Tease Out Early Dropouts

Randomized withdrawal Design

In some disease area many subjects will withdrawal early due to AE or intolerability, a design where all subjects were exposed to the study drug for a period and then randomize the patients who stayed in the trial will provide valid hypothesis testing for the drug (not necessarily estimates for the overall population)

Primary and Sensitivity Analyses

Statistician Are Not Magician

1. A trial with 50% missing data and time to event endpoint, Kaplan-Meier estimates showed a 90% cure rate. Is it credible?
2. When questioned about the estimate, clinicians will point to statisticians and common practices
3. The real issue need to be addressed is the credibility of the non-informative censoring assumption, which often is not credible

Sensitivity Should Assess Robustness to Missing

1. No one perfect analysis in dealing with missing
2. The results need to be robust to reasonable sensitivity analysis
3. Sensitivity analysis should be conservative for the comparison, not necessarily the treatment response
    - Missing as success could be more conservative than missing as failure analysis

Hepatitis B Trials

Success defined based on change of liver biopsies score is used as the primary endpoint.

1. Often these are in study missing due to concern of the risk of the liver biopsy procedure. Other lab measures like viral load and ALT are typically available
2. Often the primary analysis uses only subjects who had baseline biopsy
    - Preserves randomization but changes the population

Hepatitis B Trials

1. Missing = Failure used as the primary analysis
2. Analysis based on MAR is often encouraged. Specifically, missing is likely due to patients either feeling well or poorly and do not see added value of the procedure, and such information could be partially captured by either baseline or on treatment lab measures. Multiple imputation method could be used with a set of pre-specified predictors for the missing
3. Missing=Success analysis to cover the other extreme

Regulatory Role in Missing Data

Clearly Set the Standards for Missing Data

- Set reasonable objectives in missing data
  - For example, in a 7-14 days oral herpes trial, missing rate should not be expected to exceed 5%
  - The “soft” missing should be minimal in all trials
- In the protocol and investigator’s brochure there should be clear procedures dealing with missing data – including lost to follow-ups

Flu Guidance (for public comments)

- In these short term trials censoring subjects should be avoided. If a patient does not return for evaluation after an investigator has exhausted all reasonable means, information on the subject's status like death, description by the patient and his/her contacts on the symptoms of flu and adverse events, and general well-being of the patient should be collected and documented. This information will be very important to determine how each missing value will be regarded in the analysis. If there are still missing data despite all efforts, conservative sensitivity analysis should be conducted to demonstrate that the final conclusion is robust to the missing data.

Flu Guidance (for public comments)

- Minimizing missing data is critical in these studies because even a very low missing rate may overwhelm the small number of prophylaxis failures and make an otherwise highly significant result disappear. Investigators should be diligent in obtaining the final status of the patients either on or off the assigned treatment, either in the study or if terminated from the study.

- Subjects with diary cards that are missing for several days (i.e., less than one week) or with negative laboratory confirmation and missing their follow-up serology assessment should be considered to have missing data. Subjects with missing data in community and nursing home studies are counted as not having symptomatic laboratory-confirmed influenza. A household with no confirmed cases of influenza that has at least one contact case withdraw from the study should be defined as a household with missing data. Households with missing data are counted as not having symptomatic laboratory-confirmed influenza in the primary analysis.

Incentives for Better Follow-up

- Provide incentives for better follow-up by reasonably conservative analysis on missing data
  - Missing as success (censored) on comparator but missing as failure on the study drug
- Labels should also reflect conservative, penalized analysis instead of “neutral” analysis.
- Off-treatment follow-up data can alleviate some of the penalties

Example

	Missing (soft)	Success	Failure
Study arm	20% (10%)	60%	20%
Control	20% (10%)	40%	40%

Neutral: effect size is 60%-40%=20%

Conservative: 60%-50%=10% if missing as success on comparator arm

Conservative: 60%-40/90=16% if missing censored on comparator arm

Conclusions

- Missing is often a valid outcome that need to be interpreted
- We need to ask the right question
- We can do better in gathering information by extra effort in follow-up
- Better design is the first step.
- Incentives through reasonably conservative analysis on missing data

END

Outline

- Missing, or Not Missing - Fundamental Issues
- Toward More Informative Missing
- Regulatory Role in Missing Data
- Quality by Design
  - Design and Analysis to Fit the Objectives
  - Independent Verification
  - Limited Disclosure
  - Reducing Bias in Trial
- Resolving Missing Data Problem
  - Reduce Missing Data by Better Design, Better Data Collection, Better Efforts, Better Prioritization, and Proper Endpoint Selection
  - Collecting proper variables to aid analysis
  - Off treatment follow-up
  - What Are the Appropriate Questions?
  - What are the imputed value represents?
  - Primary and Sensitivity Analyses

Background

- Statistics is not just the analysis, it is the process from design, data generation to analysis
- Statistics is largely responsible for efficacy determination, which requires carefully consideration of the trial
- It is not just about drug approval, but quality trial will enable us to answer many questions more precisely and address more questions.
- This is based on NDA cases in anti-viral

Quality by Design

Objectives Determine Design and Analysis

- Confirmatory trials are usually for determination if a new therapy/strategy is beneficial or harmful to the patients
  - Main objectives
  - Secondary Objectives
  - Benefits should be long-term
    - Proper use of surrogate endpoints

Example

- In a HIV treatment trial for a highly treatment experienced population, the sponsor want to compare two different regimens to see which one is a better initial treatment.
- The trial will be two-arm, parallel, randomized design. Whenever a patient fails on a regimen, a new regimen will be constructed by the investigator. This strategy will be repeated over time. The trial will end at the 5^{th} year.

Example

- The proposed endpoint is the suppression of HIV viral load below detection limit of a HIV assay.
- Because suppression below assay limit is considered a validated surrogate endpoint, and a benefit on this endpoint would predict a survival benefit

- This is false
  - A subject who is suppressed on his/her 3rd regimen will likely have very limited future treatment options than a person who is still on the first regimen but begin failing.
  - Transient benefits on surrogate markers is a surrogate for clinical benefits in some settings, not necessarily in all settings

Independent Verification

- Principle in separation of power
- Lack of independent verification will lead to deterioration of the quality of trial, or at least perceived quality of trial
  - Incentives to cheat – save time and cost
  - Incentives to please – for future contracts
  - Incentives to cover up – for errors
  - Cheat due to fear of competitors cheat

Independent Verification

- Treatment Allocation
  - A single entity handling randomization, data entry, and analysis is not desirable
    - Potential for recoding the randomization code to produce favorable results
  - Need to have verifiable, authentic original randomization code, preferably maintained by a 3rd party
    - Verification of electronic documents could be particularly problematic because lower barrier for manipulations, and less traceable in general

Independent Verification

- Lab reports should be kept on site as well as in the lab for cross checking. This is especially important when lab reports are stored electronically and less traceable for modifications

- May be a centralized electronic storage could serve this purpose for the industry?
  - Data stored but not modifiable by any party
  - Data not viewable to the storage center, example, it could be password protected and encrypted by the sponsor/owner
  - Can be sent to FDA by the administrators of the storage center upon request of the sponsor/owner, but the password will be send to FDA directly by the sponsor/owner

Limiting Disclosure of Information During Trial

- Interim Analysis or adaptive designs
  - Best handled by 3rd party without link to the trial to minimize potential for information leaking to sponsor, investigators and patients
  - Keep variables for analysis and covariates to minimum. For example, analysis of efficacy and safety by sites may lead to different changes of enrollment in different sites

Limiting Public Disclosure of Information

- Public disclosure of information may affect trial results
  - Sometimes done for investors and public relationships
- May make enrollment more difficult, may change the types of patients enrolled, and may introduce biases
  - Discouraged
  - May require modification of analysis to handle the data pre- and post- release data differently

Reducing Biases in Clinical Trials

- Open-label trials lead to potential bias
  - Lead to investigator and patient change in the assessment of the outcome, especially subjective measures
  - Lead to different behaviors for different arms by both the investigator and the patients
    - Investigator may provide additional supplemental care to the patients who are taking a regimen which is considered less efficacious
    - Patients may failed on the control regimen on purpose to get the new drug

Reducing Biases in Clinical Trials

- Trials should be blinded whenever possible
  - This is better recognized for confirmatory trials
  - Less so for the exploratory trials or Phase IV trials
    - Biased outcome in exploratory trials will lead to incorrect and costly decisions, either for unnecessary further development of the drug, or lost opportunity in development

Reducing Biases in Clinical Trials

- Blinded trials could also be biased
  - Blinding can be partially defeated by adverse events or responses in certain lab markers
  - Sloppiness in trial conduct, for example, failing to adhere to the assigned regimen or schedule, poor instruments for measurements could make two treatment arms similar even when they are truly different
    - Reduce the effect size in a superiority trial
    - Make an inferior new drug appear similar to an established regimen

Reducing Biases in Clinical Trials

- Adherence to the pre-specified analysis plan
  - Journal publications often fail to mention which analysis is pre-planned, which is exploratory
    - Example, one article showed an odds ratio of 5.03, after adjusting for 6 covariates in a trial with 42 patients. The journal failed to show that without adjusting the covariates the odds ratio is 3.7, and if the baseline adjustment was pre-specified
  - Failure to mention or verify if the displayed analysis is pre-specified is a cause for the misleading results. This adds upon the publication bias due to only publishing good results
  - Significantly reduces the usefulness and credibility for regulatory use

Handling Biases in Clinical Trials

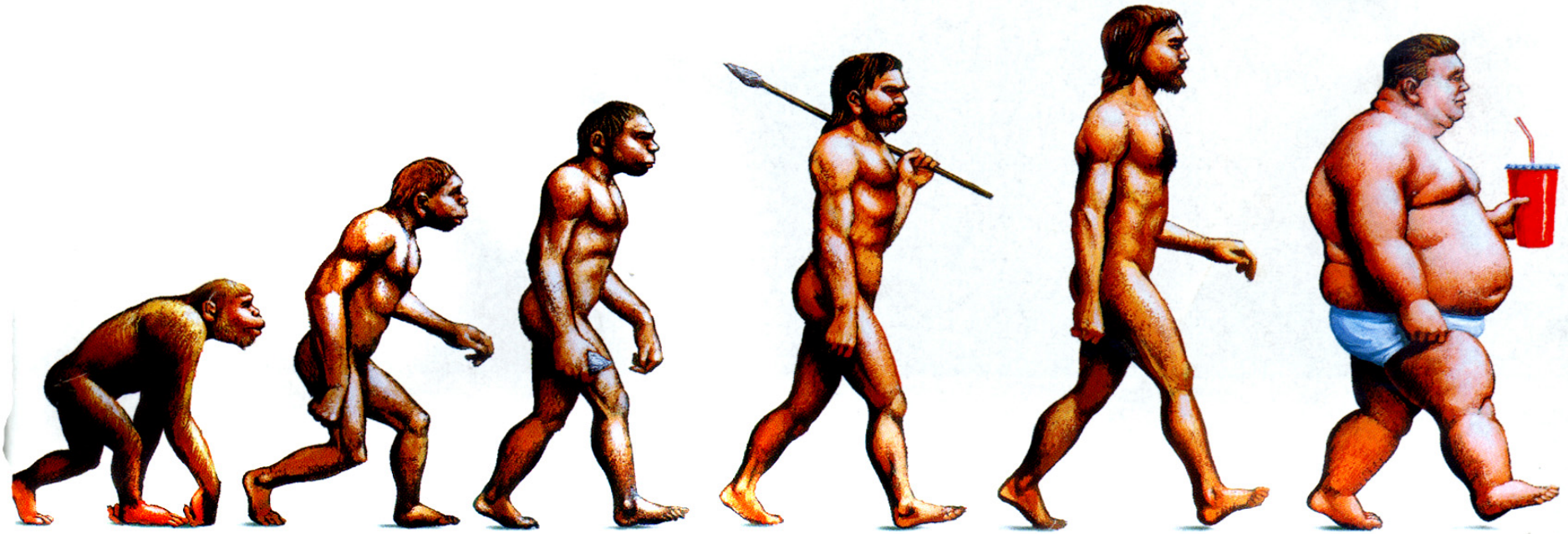
- Anticipate potential scenarios and collecting data as potential markers
  - If subjects on a perceived less efficacious regimen has the potential of being less compliant in order to have early access to open label new drug, then measures like pill counts and measure of drug concentration may provide some indication. Very early failures are possible candidates.
  - In statistical analysis their failure status need to be handled in a conservative way, for example, censored.

Handling Biases in Clinical Trials

- Biases can not be corrected through alpha level adjustment
  - Any penalty in alpha level will not address the issue. With large sample size, the biases will become efficacy no matter how small your alpha is.
  - Confirmation from another trial will not address the bias issue. Two biased trials confirming each other is worse than a single biases trial

Missing Data

The shape of things to come



- “Effect of ganciclovir therapy on hearing in symptomatic congenital cytomegalovirus disease involving the central nervous system: a randomized, controlled trial” and published in Journal of Pediatrics, 2003; 143:16-25