

【企画セッション】欠測のあるデータの解析のためのSASプログラムの紹介

～データ発生・DIAマクロとプロシジャの進展～

(2) シミュレーションデータ発生方法

横山雄一^{1,2}, 大浦智紀^{1,3}, 土居正明^{1,4}

(¹日本製薬工業協会 医薬品評価委員会 データサイエンス部会 タスクフォース4
欠測のあるデータの解析検討チーム,

² 持田製薬株式会社, ³ 日本イーライリリー株式会社, ⁴ 東レ株式会社)

【Planning session】 Introduction of the SAS programs to analyze the missing data

(2) The method of simulation data generation

Yuichi Yokoyama^{1,2}, Tomonori Oura^{1,3}, Masaaki Doi^{1,4}

(¹ The team for statistical analysis of data analysis with missing data, task force 4,
data science expert committee, drug evaluation committee,

Japan Pharmaceutical Manufacturers Association,

² Mochida Pharmaceutical Co., Ltd., ³ Eli Lilly Japan K. K., ⁴ Toray Industries, Inc.)



要旨:

文献データから、欠測のある連続量経時データを発生させる方法について、SASプログラムとともに紹介する。

キーワード:

文献データ, シミュレーションデータ, データ発生,
SASプログラム

発表構成

1. 背景
2. 文献データ
3. シミュレーションデータの設定
4. シミュレーションデータ発生方法
5. シミュレーションにおける留意点
6. まとめと今後の課題



1. 背景

背景 (1/2)

- ・ 応答変数が経時的に測定される臨床試験において、データの欠測はよく起こる.
- ・ 欠測のあるデータに対しては、様々な解析手法があり、試験や解析の目的に合わせた使い分けが必要.
- ・ 計画時に試験及び薬剤の特徴を捉えたシミュレーションデータを発生させ、目的に合わせた検討を行うことは有用.
- ・ 本セッションでは、応答変数は連続量のみを扱う.

【シミュレーションの目的】

- ・ 主解析の評価 (α エラー, 検出力, 推定精度)
- ・ 感度分析の選択
- ・ 詳細な仕様の特定 (推定方法, 分散共分散構造, 等)

背景(2/2)

- これまでに検討してきたシミュレーションデータ
 - ① うつ病データ
 - SASユーザー総会 2014の土居ら(2014)で検討.
 - ② 慢性疼痛データ
 - 日本製薬工業協会シンポジウム「臨床試験の欠測データの取り扱いに関する最近の展開と今後の課題について
—統計手法・estimandと架空の事例に対する流れの整理—」の横山ら(2015)で検討.

【本発表の目的】

- 文献データから、シミュレーションデータを発生させる方法を紹介する.



2. 文献データ

文献データ (Eli Lilly, 2004) (1/2)

- 試験の目的
 - 大うつ病性障害患者を対象に, 9週間投与後の試験薬群のプラセボ群に対する優越性を検証する (Phase 3).
- 被験者構成
 - 実薬群 (n=128)
 - 完了 (60.9%), 中止: AE (12.5%), 効果不十分 (5.5%)
 - プラセボ群 (n=139)
 - 完了 (64.7%), 中止: AE (4.3%), 効果不十分 (13.7%)
- Estimand 3 (Efficacy)
 - 計画された時点での, 「全症例が治療を完了した」と仮定した場合の応答変数の改善の群間差.

計画時: n=120/群

文献データ (Eli Lilly, 2004) (2/2)

- 統計解析

- HAM-Dの変化量について, MMRMにより, 実薬群とプラセボ群を比較する.
 - 主要な評価時点は最終時点 (Week 9).

- 解析結果

Week	N		LSMean		LSMean Change(SE)		p値 (vs.プラセボ群)
	実薬群	プラセボ群	実薬群	プラセボ群	実薬群	プラセボ群	
0	123	136	20.28※	20.49※			
1	123	136	17.58	17.83	-2.89(0.38)	-2.64(0.36)	0.601
2	109	129	14.93	16.05	-5.54(0.48)	-4.43(0.45)	0.071
3	108	122	13.65	14.42	-6.82(0.55)	-6.06(0.52)	0.287
5	98	111	11.89	13.27	-8.58(0.66)	-7.20(0.62)	0.116
7	89	97	10.34	12.78	10.14(0.69)	-7.69(0.65)	0.008
9	81	90	10.01	12.18	10.46(0.71)	-8.29(0.67)	0.024

※ Mean Baseline

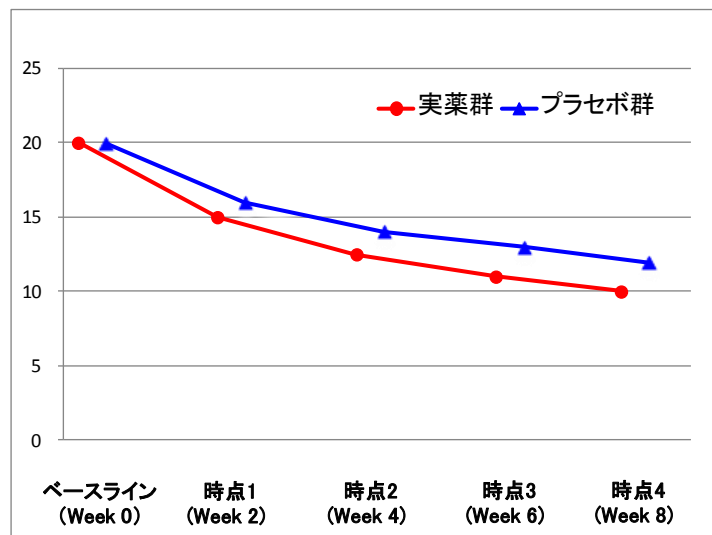


3. シミュレーションデータの設定

シミュレーションデータの設定 (1/2)

完全データ

被験者数: 120例/群



完全データの平均構造

時点間の相関 ※文献に記載なし

	ベースライン	時点1	時点2	時点3	時点4
ベースライン	1.00	0.70	0.60	0.50	0.40
時点1	-	1.00	0.70	0.60	0.50
時点2	-	-	1.00	0.70	0.60
時点3	-	-	-	1.00	0.70
時点4	-	-	-	-	1.00

各時点の測定値の平均 (SD)

測定値の平均 (SD)	ベースライン	時点1	時点2	時点3	時点4
実薬群	20.0 (4.0)	15.0 (5.0)	12.5 (6.0)	11.0 (7.0)	10.0 (8.0)
プラセボ群	20.0 (4.0)	16.0 (5.0)	14.0 (6.0)	13.0 (7.0)	12.0 (8.0)

シミュレーションデータの設定 (2/2)

欠測確率の関数 ※文献に記載なし

単調な欠測を仮定

- ・ MCARとMNARの欠測メカニズムを1:1の割合で混合
→ 中止理由によって、欠測メカニズムが異なることを考慮した。

MCAR

時点ごとの欠測確率の目標値

群	時点1	時点2	時点3	時点4
実薬群	20%	10%	10%	10%
プラセボ群	10%	10%	10%	10%

MNAR

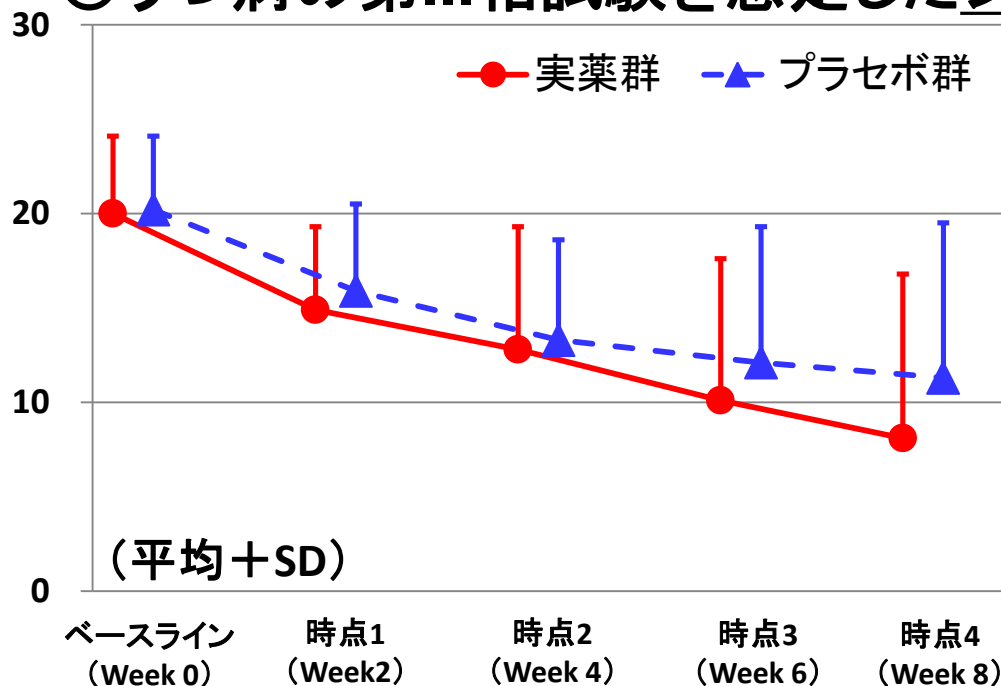
- ・ 群ごとで同じ関数形
→ 応答変数 Y_{t-1} , Y_t の推移が群ごとで異なる
→ 群ごとに異なる欠測確率

Y_t : 時点 t の測定値
 p_t : 時点 $t-1$ で観測された症例が
 時点 t で欠測する確率

$$\text{logit}(p_t) = -9.5 + 0.2 \cdot Y_{t-1} + 0.2 \cdot Y_t$$

解析対象データ

◎うつ病の第III相試験を想定したシミュレーションデータ



- 主要評価項目: HAM-D
→ **スコア低下: 改善**
- 実薬群 vs プラセボ群
・1群120例(ベースライン)

	ベースライン		時点1		時点2		時点3		時点4	
	例数	平均 (SD)	例数	平均 (SD)	例数	平均 (SD)	例数	平均 (SD)	例数	平均 (SD)
実薬群	120	20.0 (4.1)	100	14.9 (4.4)	92	12.8 (6.5)	86	10.1 (7.5)	83	8.1 (8.7)
プラセボ群	120	20.2 (3.9)	106	15.9 (4.6)	100	13.3 (5.3)	94	12.1 (7.2)	88	11.3 (8.2)

解析結果

- シミュレーションデータ

時点	N		LSMean		LSMean Change(SE)		p値 (vs.プラセボ群)
	実薬群	プラセボ群	実薬群	プラセボ群	実薬群	プラセボ群	
1 (Week 2)	100	106	14.97	15.77	-4.88(0.33)	-4.08(0.36)	0.084
2 (Week 4)	92	100	12.87	13.29	-6.98(0.50)	-6.56(0.48)	0.539
3 (Week 6)	86	94	10.61	12.13	-9.24(0.68)	-7.72(0.65)	0.108
4 (Week 8)	83	88	8.69	11.56	-11.16(0.85)	-8.29(0.82)	0.016

- 文献データ(参考)

Week	N		LSMean		LSMean Change(SE)		p値 (vs.プラセボ群)
	実薬群	プラセボ群	実薬群	プラセボ群	実薬群	プラセボ群	
1	123	136	17.58	17.83	-2.89(0.38)	-2.64(0.36)	0.601
2	109	129	14.93	16.05	-5.54(0.48)	-4.43(0.45)	0.071
3	108	122	13.65	14.42	-6.82(0.55)	-6.06(0.52)	0.287
5	98	111	11.89	13.27	-8.58(0.66)	-7.20(0.62)	0.116
7	89	97	10.34	12.78	-10.14(0.69)	-7.69(0.65)	0.008
9	81	90	10.01	12.18	-10.46(0.71)	-8.29(0.67)	0.024



4. シミュレーションデータ発生方法

シミュレーションデータの発生の流れ

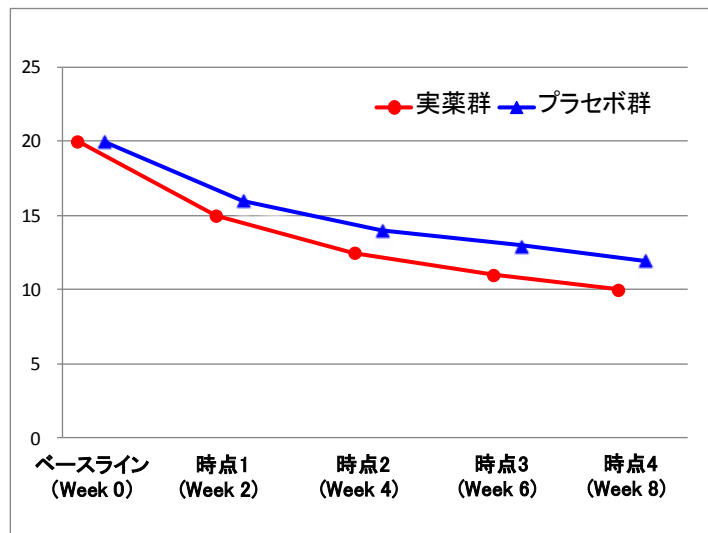


評価時点の設定

- ・ 試験及び薬剤の特徴を捉えた必要最低限の時点を設定することが好ましい。

【シミュレーションデータの特徴】

- ・ 実薬群：中盤以降での薬効発揮，有害事象での早期中止
- ・ プラセボ群：緩やかなスコアの減少，効果不十分での中止



完全データの平均構造

- ・ うつ病の第Ⅲ相試験を想定
- ・ 応答変数：HAM-D (重症度スコア)
- ・ 群：2群 (実薬群, プラセボ群)
- ・ 被験者数：120例/群

完全データの発生 (1/3)

1. 群・時点毎に平均及びSDを設定する.
2. 被験者の影響を考慮する(2種類の方法).

方法1. 変量効果(変量切片)を指定する.

$$Y_i = X_i\beta + \mathbf{1}_n b_i + \epsilon_i$$
$$b_i \sim N(0, \nu^2), \quad \epsilon_i \sim N(0, \sigma^2 \mathbf{I}_n)$$

- ・ 変量効果は1変量(被験者)のみ
- ・ 誤差は時点間で独立と想定してデータを発生

方法2. 時点間の相関を指定する.

$$Y_i = X_i\beta + \epsilon_i$$
$$\epsilon_i \sim N(0, \Sigma) \quad \Sigma \text{ は, 誤差の分散共分散行列}$$

- ・ 変量効果なし
- ・ 誤差の相関構造(相関行列)を指定してデータを発生

完全データの発生 (2/3)

- SDと相関行列から分散共分散行列を算出する。

$$r_{ij} = \frac{\sigma_{ij}}{\sigma_i \cdot \sigma_j}$$

各時点の測定値の平均 (SD)

測定値の平均 (SD)	ベースライン	時点1	時点2	時点3	時点4
実薬群	20.0 (4.0)	15.0 (5.0)	12.5 (6.0)	11.0 (7.0)	10.0 (8.0)
プラセボ群	20.0 (4.0)	16.0 (5.0)	14.0 (6.0)	13.0 (7.0)	12.0 (8.0)

Toeplitz Structure

誤差の相関構造 (相関行列)

$$R = \begin{pmatrix} 1.00 & 0.70 & 0.60 & 0.50 & 0.40 \\ 0.70 & 1.00 & 0.70 & 0.60 & 0.50 \\ 0.60 & 0.70 & 1.00 & 0.70 & 0.60 \\ 0.50 & 0.60 & 0.70 & 1.00 & 0.70 \\ 0.40 & 0.50 & 0.60 & 0.70 & 1.00 \end{pmatrix}$$

分散共分散行列

$$\Sigma = \begin{pmatrix} 16.0 & 14.0 & 14.4 & 14.0 & 12.8 \\ 14.0 & 25.0 & 21.0 & 21.0 & 20.0 \\ 14.4 & 21.0 & 36.0 & 29.4 & 28.8 \\ 14.0 & 21.0 & 29.4 & 49.0 & 39.2 \\ 12.8 & 20.0 & 28.8 & 39.2 & 64.0 \end{pmatrix}$$

完全データの発生 (3/3)

- ・ 設定した平均ベクトル及び分散共分散行列をもとに、多変量正規乱数を発生させる。

```
PROC IML;
```

```
mean1={ 20.0 15.0 12.5 11.0 10.0 };  
mean2={ 20.0 16.0 14.0 13.0 12.0 };
```

平均ベクトル

```
cov={ 16.0 14.0 14.4 14.0 12.8,  
      14.0 25.0 21.0 21.0 20.0,  
      14.4 21.0 36.0 29.4 28.8,  
      14.0 21.0 29.4 49.0 39.2,  
      12.8 20.0 28.8 39.2 64.0 };
```

分散共分散行列

```
CALL RANDSEED(12345);  
rv1=RandNormal(100, mean1, cov);  
rv2=RandNormal(100, mean2, cov);
```

多変量正規乱数の発生

```
CREATE d1 FROM rv1[colname={'y0' 'y1' 'y2' 'y3' 'y4'}];  
APPEND FROM rv1; APPEND FROM rv2;
```

SASデータセットの作成

```
quit;
```

Sample SAS code

欠測データの特定 (1/3)

1. 群・時点毎に欠測確率の目標値を設定する.
 2. 欠測メカニズムを設定する(2種類の方法).
 - 方法1. 1つの欠測メカニズムを仮定する.
 - 方法2. 複数の欠測メカニズムの混合を仮定する.
MCARとMNARの欠測メカニズムを1:1の割合で混合
- 臨床試験における代表的な中止理由
 - 有害事象
 - 効果不十分
 - 被験者都合
- ⇒ 被験者毎に欠測メカニズムが異なるかもしれず、群によってその割合が異なるかもしれない。

欠測データの特定 (2/3)

- 欠測メカニズムごとに対応する欠測確率の関数を設定する。

MCAR

時点ごとの欠測確率の目標値

群	時点1	時点2	時点3	時点4
実薬群	20%	10%	10%	10%
プラセボ群	10%	10%	10%	10%

MNAR

1時点前の測定値とその時点の測定値が
大きいほど欠測しやすい状況

Y_t : 時点 t の測定値
 p_t : 時点 $t-1$ で観測された症例が
時点 t で欠測する確率

$$\text{logit}(p_t) = -9.5 + 0.2 \cdot Y_{t-1} + 0.2 \cdot Y_t$$

実際には...各係数は、最終時点の欠測確率が
文献データの値に近づくように試行錯誤で設定

欠測データの特定 (3/3)

```
data d3; set d2;
  call streaminit(123);
  array y{5} y0-y4;
  array p{4} p1-p4;
  array m{4} m1-m4;

  if TYPE="MCAR" then do;
    do t=1 to 4;
      if trt=1 and t eq 1 then p{t}=0.2;
      if trt=1 and t ne 1 then p{t}=0.1;
      if trt=2          then p{t}=0.1;
      m{t}=rand('bernoulli',p{t});
    end; end;

  if TYPE="MNAR" then do;
    do t=1 to 4;
      p{t}=1 / ( 1 + exp( - (-9.5 + 0.2*y{t} + 0.2*y{t+1}) ));
      m{t}=rand('bernoulli',p{i});
    end; end;

run;
```

Sample SAS code

y0 : ベースラインの測定値
y1-y4 : 時点1~4の測定値
p1-p4 : 時点1~4で欠測する(条件付き)確率
m1-m4 : 時点1~4の欠測識別変数(1:欠測)



5. シミュレーションにおける留意点

シミュレーションにおける留意点(1/2)

- 参考データの結果を一般化し過ぎない。
 - 1シナリオのシミュレーション結果が完全に再現されることは稀。
 - 想定される複数のパターンのシミュレーションを実施する。
 - 試験及び薬剤の特徴を考慮し、シンプルな設定とする。
 - 実現値に対して詳細に設定を行ってしまうと、オーバーフィッティングによって一般性が下がってしまうことがある。

【文献データの利用】

- 「記載のある数値」「記載のない数値」を区別する。
 - ① 相関構造
 - ② 時点ごとの中止理由, 中止確率
- 「記載のある数値」の中でも、再現性が高そうなものと低そうなものを区別する。

シミュレーションにおける留意点(2/2)

- ・ 以下の点などを考慮して、試験ごとに十分な検討が必要。

【欠測メカニズム】

- ・ MARを仮定して本当に大丈夫か？
 - ・ MNARを仮定したシミュレーションも合わせて行うべき。

【データの分布】

- ・ 外れ値がある場合
- ・ 欠測データが観測データと大きく異なる分布の場合
- ・ ベースラインと時点の交互作用がある場合

【欠測確率の関数】

- ・ 欠測確率と応答変数の関係
- ・ 補助変数がある場合



6. まとめと今後の課題

まとめと今後の課題(1/2)

【まとめ】

- ・ 文献データに基づく、シミュレーションデータの設定
- ・ シミュレーションデータ発生方法
 - ・ 完全データの発生, 欠測データの特定
- ・ シミュレーションにおける留意点

※計画段階ではシミュレーション等を通じて、試験及び薬剤の特徴を踏まえた十分な検討をすることが必要。
ただし、結果の一般化には注意が必要。

まとめと今後の課題(2/2)

【今後の課題】

- 欠測確率の関数の精査
 - SM型の関数だけでなくPMM型, SPM型も検討する.
 - 群ごとの中止理由の違いを, より適切に組み込む.
 - 選択基準等による中止確率の変動を想定し, 適切な脱落確率モデルを複数設定する.
 - (投与後の)補助変数に脱落確率が依存する場合

参考文献

1. Ayele, B. T., Lipkovich, I., Molenberghs, G. and Mallinckrodt, C. H. (2014). A multiple-imputation-based approach to sensitivity analysis and effectiveness assessments in longitudinal clinical trials. *Journal of Biopharmaceutical Statistics*. **24**, 211-228.
2. Detke, M. J., Lu, Y., Goldstein, D. J., McNamara, R. K. and Demitrack, M. A. (2002). Duloxetine 60 mg once daily dosing versus placebo in the acute treatment of major depression. *Journal of Psychiatric Research*. **36**, 383-390.
3. 土居正明 ほか. (2014). 欠測のあるデータに対する総合的な感度分析と主解析の選択. SASユーザー総会論文集.
4. Eli Lilly (2004). Duloxetine Once-Daily Dosing Versus Placebo in the Acute Treatment of Major Depression.
<http://www.lillytrials.com/results/Cymbalta.pdf>
5. Siddiqui, O., Hung, H. M., and O'Neill, R. (2009). MMRM vs. LOCF: a comprehensive comparison based on simulation study and 25 NDA datasets. *Journal of Biopharmaceutical Statistics*. **19**, 227-246.
6. Wicklin, R. (2013). *Simulating Data with SAS*. SAS Press.
7. Wicklin, R. (2010). *Statistical Programming with SAS/IML Software*. SAS Press.
8. 横山雄一, 横溝孝明, 大浦智紀, 大江基貴. (2015). 【日本製薬工業協会シンポジウム】臨床試験の欠測データの取り扱いに関する最近の展開と今後の課題について -統計手法・estimandと架空の事例に対する流れの整理- (7)架空の事例2 (主解析の選択・例数設計・データの発生方法)
http://www.jpma.or.jp/information/evaluation/symposium/pdf/20150212/20150212_7.pdf



APPENDIX

シミュレーションデータ発生

```
PROC IML;
  mean1={ 20.0 15.0 12.5 11.0 10.0 };
  mean2={ 20.0 16.0 14.0 13.0 12.0 };

  cov={ 16.0 14.0 14.4 14.0 12.8,
        14.0 25.0 21.0 21.0 20.0,
        14.4 21.0 36.0 29.4 28.8,
        14.0 21.0 29.4 49.0 39.2,
        12.8 20.0 28.8 39.2 64.0 };

  CALL RANDSEED(12345);
  rv1=RandNormal(120, mean1, cov);
  rv2=RandNormal(120, mean2, cov);

  trt1=J(120, 1, 1); trt2=J(120, 1, 2);
  rv1b=trt1||pat||rv1; rv2b=trt2||pat||rv2;
  CREATE d1 FROM rv1b[colname={'trt' 'y0' 'y1' 'y2' 'y3' 'y4'}];
  APPEND FROM rv1b; APPEND FROM rv2b;
quit;

data d0;
  format id trt;
  do nobs=1 to 120;
    trt=1; id=1000+nobs; uni=ranuni(4649); output;
    trt=2; id=2000+nobs; uni=ranuni(5963); output;
  end;
  drop nobs;
run;

data d2; merge d0 d1;
  if 0.0<=uni<0.5 then uni_flg=1;
  else if 0.5<=uni<1.0 then uni_flg=2;
run;
```

Sample SAS code

```
data d3; set d2;
  call streaminit(123);
  array y{5} y0-y4;
  array p{4} p1-p4;
  array m{4} m1-m4;

  if uni_flg=1 then do; *- MCAR;
    do t=1 to 4;
      if trt=1 and t eq 1 then p{t}=0.2;
      if trt=1 and t ne 1 then p{t}=0.1;
      if trt=2 then p{t}=0.1;
      m{t}=rand('bernoulli',p{t});
    end; end;

  if uni_flg=2 then do; *- MNAR;
    do t=1 to 4;
      p{t}=1 / ( 1 + exp( - (-9.5 + 0.2*y{t} + 0.2*y{t+1})));
      m{t}=rand('bernoulli',p{t});
    end; end;

  if m1=1 then do; y1=.; y2=.; y3=.; y4=.; end;
  if m2=1 then do; y2=.; y3=.; y4=.; end;
  if m3=1 then do; y3=.; y4=.; end;
  if m4=1 then do; y4=.; end;

  keep id trt y0-y4;
run;
```