

## Control for Selection Bias Related to the Choice of Therapy

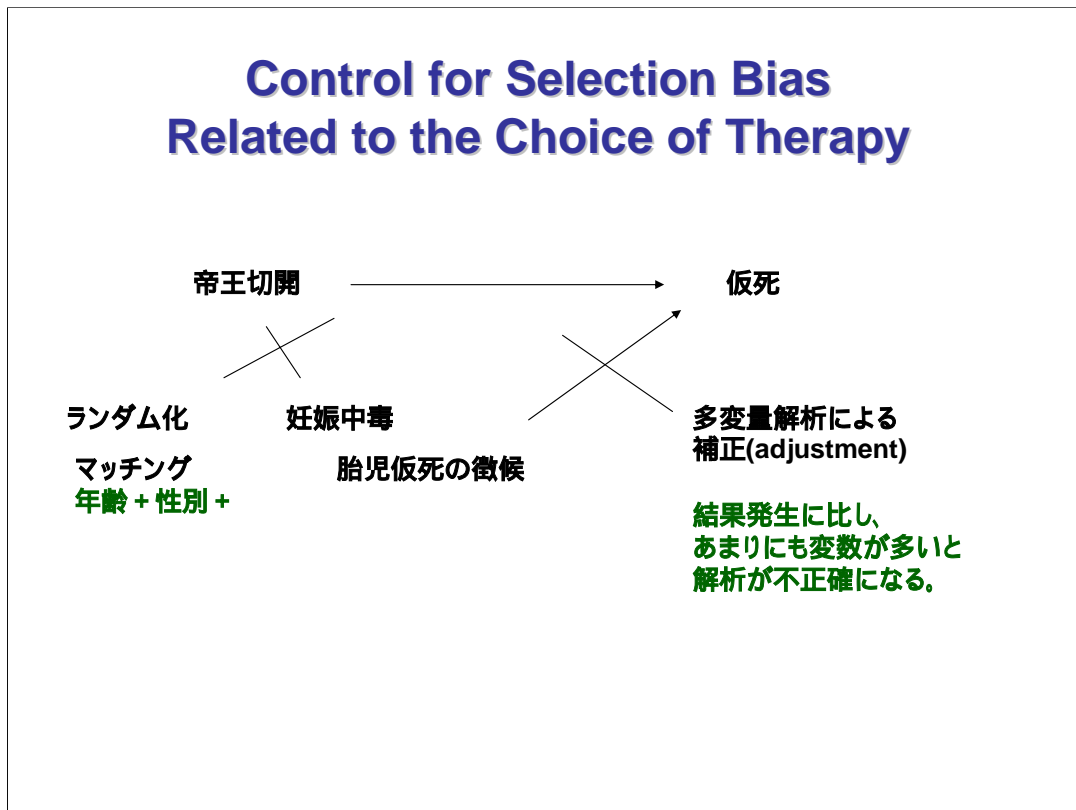


図1. 帝王切開は古くから行なわれてきた医療である。胎児仮死の徴候があれば、産科医は帝王切開に踏み切る。しかし、胎児仮死の徴候があれば出生時仮死のリスクも高くなる。そのため、一見帝王切開が仮死に結びつく錯覚におちいるが、さまざまな交絡因子が関与するため、そのような結論をくだすわけにはいかない。このような場合、仮死のあった児(ケース)と妊娠中毒症などの程度をマッチングさせて仮死の無かった児(コントロール)を選ぶマッチングという手法がある。一方、多変量解析により交絡因子の影響を補正する方法もある。このようにして純粋な帝王切開の仮死に及ぼす影響について知ることができる。

Am J Epidemiol. 2003 Aug 1;158(3):280-7. [Related Articles, Links](#)

**Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders.**

- The aim of this study was to use Monte Carlo simulations to compare logistic regression with propensity scores in terms of bias, precision, empirical coverage probability, empirical power, and robustness when the number of events is low relative to the number of confounders. **The authors simulated a cohort study and performed 252,480 trials.** In the logistic regression, the bias decreased as the number of events per confounder increased. In the propensity score, the bias decreased as the strength of the association of the exposure with the outcome increased. **Propensity scores produced estimates that were less biased, more robust, and more precise than the logistic regression estimates when there were seven or fewer events per confounder.** The logistic regression empirical coverage probability increased as the number of events per confounder increased. The propensity score empirical coverage probability decreased after eight or more events per confounder. **Overall, the propensity score exhibited more empirical power than logistic regression. Propensity scores are a good alternative to control for imbalances when there are seven or fewer events per confounder; however, empirical power could range from 35% to 60%. Logistic regression is the technique of choice when there are at least eight events per confounder.**

図2 . Am J Epidemiol. 2003 Aug 1;158(3):280-7.

Am J Epidemiolにpropensity score の適応例が示された。すなわち、結果発生数に変数の数の7倍以下である場合、propensity score を用いた方が的確な回答を得られるというのである。

## Control for Selection Bias Related to the Choice of Therapy

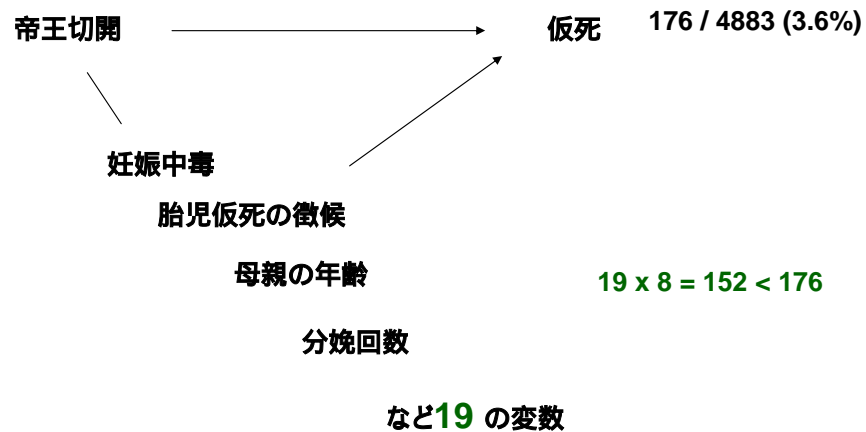


図3 . しかし、現実問題として仮死はそう多いものではない。今回の例では、4883人の出産を調査し、176人(3.6%)に仮死をみた。仮死に影響するのは妊娠中毒だけではなく、母親の年齢、分娩回数など多くの変数がある。具体的には結果発生(この例の場合仮死)数の数が変数 $\times 8$ より大きいとき、多変量解析で十分である。

## Control for Selection Bias Related to the Choice of Therapy

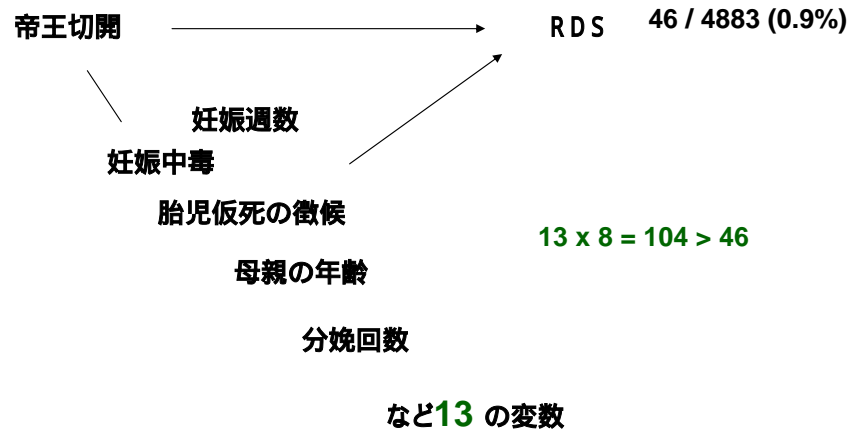


図4 . 一方、結果を呼吸窮迫症候群に設定した場合、結果は僅か46人(0.9%) にしか認められなかった。結果に影響し得る変数を13選んだとすると、 $13 \times 8 = 104$  で46人より多くなる。このような場合、propensity scoreの方が精度が高くなる。実際、このように結果発生が少ない状況で多変量解析を行なうと、個々の因子の有意性がブレやすい。

```
. tab caiser emergentcaiser
```

bunben== 3.0000	bunben== 0	4.0000 1	Total
0	3,870	430	4,300
1	583	0	583
Total	4,453	430	4,883

図5 . 実際のお産のデータである。全部で4883例のデータがここにある。583人が予定帝王切開を、430人が緊急帝王切開を受けて生まれている。

```
. logistic allcaiser chudoku age bunkai shu twin hasui cpd zenchi iugr kaisen toui nojomyaku
```

```
Logistic regression          Number of obs   =    4879  
                             LR chi2(12)         =   1728.56  
                             Prob > chi2          =    0.0000  
Log likelihood = -1627.8735   Pseudo R2       =    0.3468
```

allcaiser	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
chudoku	4.01112	.6658452	8.37	0.000	2.897119	5.553477
age	1.093227	.0112102	8.69	0.000	1.071475	1.115421
bunkai	.889435	.051485	-2.02	0.043	.79404	.9962907
shu	.7828724	.0159203	-12.04	0.000	.7522829	.8147058
twin	1.612728	.1349424	5.71	0.000	1.368794	1.900133
hasui	.5104511	.060805	-5.65	0.000	.4041652	.6446877
cpd	44.87512	11.29705	15.11	0.000	27.39812	73.50055
zenchi	59.75519	37.00279	6.61	0.000	17.75325	201.1284
iugr	6.258763	1.566963	7.33	0.000	3.83159	10.22346
kaisen	1.88025	.4955546	2.40	0.017	1.121695	3.151784
toui	.0659607	.0101835	-17.61	0.000	.0487384	.0892689
nojomyaku	4.53996	.6218002	11.05	0.000	3.471124	5.937915

図6 . まず緊急、予定帝王切開が選ばれる因子を多ロジスティック解析で検討する。

```

Stata Results
. logit

Logit estimates                               Number of obs   =      4879
LR chi2(12)                                   =     1728.56
Prob > chi2                                    =       0.0000
Pseudo R2                                     =       0.3468

Log likelihood = -1627.8735

+-----+-----+-----+-----+-----+-----+
| allcaiser |      Coef. |   Std. Err. |      z | P>|z| | [95% Conf. Interval] |
+-----+-----+-----+-----+-----+-----+
| chudoku   |    1.38907 |   .1659998   |    8.37 | 0.000 |    1.063717   1.714424 |
| age       |   -.0891338 |   .0102542   |   -8.69 | 0.000 |   -.0690358  -.1092317 |
| bunkai    |  -.1171689 |   .0578851   |   -2.02 | 0.043 |  -.2306215  -.0037162 |
| shu       |  -.2447855 |   .0203357   |  -12.04 | 0.000 |  -.2846429  -.2049282 |
| twin      |   .4779271 |   .0836734   |    5.71 | 0.000 |   .3139302   .641924 |
| hasui     |  -.6724604 |   .1191201   |   -5.65 | 0.000 |  -.9059315  -.4389892 |
| cpd       |   3.803884 |   .251744    |   15.11 | 0.000 |   3.310474   4.297293 |
| zenchi    |   4.090256 |   .6192397   |    6.61 | 0.000 |   2.876569   5.303944 |
| iugr      |   1.833983 |   .2503631   |    7.33 | 0.000 |   1.34328    2.324685 |
| kaisen    |   .6314049 |   .2635578   |    2.40 | 0.017 |  -.1148411   1.147969 |
| toui      |  -2.718696 |   .1543871   |  -17.61 | 0.000 |  -3.021289  -2.416103 |
| nojomyaku |   1.512918 |   .1369616   |   11.05 | 0.000 |   1.244478   1.781358 |
| _cons     |    6.1899  |   .8419788   |    7.35 | 0.000 |   4.539652   7.840148 |
+-----+-----+-----+-----+-----+

. gen p=1.38907*chudoku+.0891338*age-.1171689*bunkai-.2447855*shu+.4779271*twin-.6724604*hasui+3.80
> cpd+4.090256*zenchi+1.833983*iugr+.6314049*kaisen-2.718696*toui+1.512918*nojomyaku+6.1899
<4 missing values generated>

. gen PS = exp(p)/(exp(p)+1)
<4 missing values generated>

. sum PS

+-----+-----+-----+-----+-----+
| Variable |      Obs |      Mean |   Std. Dev. |      Min |      Max |
+-----+-----+-----+-----+-----+
| PS       |      4879 |   .2076246 |   .254978    |   .0042576 |   .9990798 |
+-----+-----+-----+-----+-----+

```

図7 . そしてpropensity score (PS) を図のような要領で算出する。

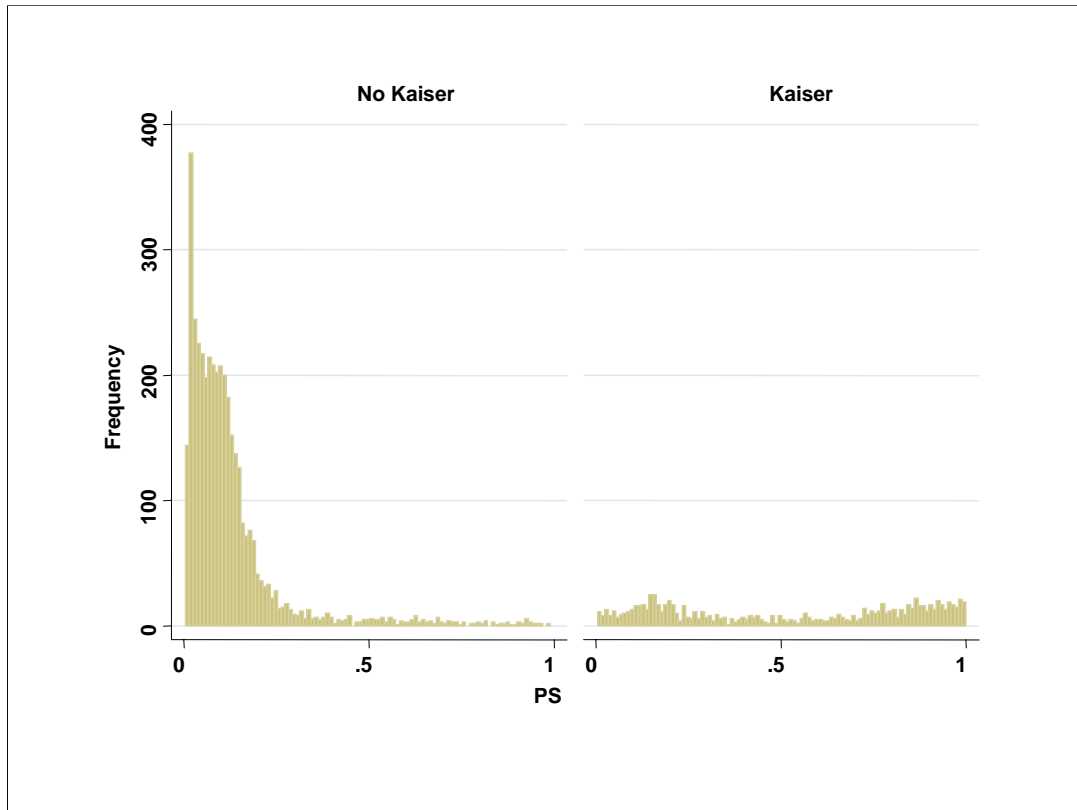


図8 . 帝王切開を行なわれた対象のpropensity scoreは0から1まで広く分布している。一方、帝王切開を施行されなかった対象では、低いpropensity scoreに集中している。



```
. tab psquantile allcaiser, row
```

Key		allcaiser		Total
	frequency	0	1	
psquantile				
1	937 96.10	38 3.90	975 100.00	
2	936 95.90	40 4.10	976 100.00	
3	899 92.49	73 7.51	972 100.00	
4	781 79.69	199 20.31	980 100.00	
5	313 32.07	663 67.93	976 100.00	
Total	3,866 79.24	1,013 20.76	4,879 100.00	

図9 . Propensity score を5段階に分け、帝王切開の行なわれた頻度を調べてみた。Propensity score が大きくなるにつれて帝王切開が選択される割合が高くなっているのが理解できる。

	BT	BU	BV	BW	BX	BY	BZ	CA	CB	CC
1	single2	md	dd	predlac	difflac	freestylebi	rdbi	allcaiser	PS	pair
2	1	0	0	3.882848	2.297152	1	0	0	0.004258	
3	1	0	0	3.882847	1.657153	0	0	0	0.004861	
4	1	0	0	3.882844	-0.17284	0	0	0	0.005432	
5	1	0	0	3.882846	1.337154	0	0	0	0.005432	
6	1	0	0	3.882849	4.167151	0	0	0	0.005432	
7	1	0	0	3.882848	2.817152	0	0	0	0.005556	
8	1	0	0	3.882846	-0.05285	0	0	0	0.005556	
9	1	0	0	3.882849	0.327152	0	0	0	0.005935	
10	1	0	0	3.882848	1.557152	0	0	0	0.00637	1
11	1	0	0	3.882847	0.097153	0	0	1	0.00645	1
12	1	0	0	3.882844	-0.12284	0	0	0	0.006633	
13	1	0	0	3.882844	-0.83284	0	0	0	0.006633	
14	1	0	0	3.882845	-0.62285	0	0	0	0.006769	
15	1	0	0	3.882846	0.127154	0	0	0	0.006928	
16	1	0	0	3.882847	1.017154	0	0	0	0.007246	
17	1	0	0	3.882845	-0.60285	0	0	0	0.007246	
18	1	0	0	3.882847	-0.05285	0	0	0	0.007246	
19	1	0	0	3.882846	1.127154	0	0	0	0.007604	
20	1	0	0	3.882845	-0.51284	0	0	0	0.007917	
21	1	0	0	3.882846	-0.87285	0	0	0	0.007917	
22	1	0	0	3.882846	2.527153	0	0	0	0.008269	
23	1	0	0	3.882849	2.387151	0	0	0	0.008307	
24	1	0	0	3.882848	1.267153	0	0	0	0.008307	2
25	1	0	0	3.88285	3.25715	0	1	1	0.008457	2
26	1	0	0	3.882847	2.047153	0	0	0	0.008457	
27	1	0	0	3.882849	2.987151	1	0	0	0.008457	
28	1	0	0	3.882845	-0.47285	1	0	0	0.008457	

図10. Propensity score (PS) と帝王切開 (allcaiser) をソートしたEXCEL表である。10番目の対象は帝王切開を施行されているが、9番目の対象は帝王切開を施行されていないので、これをマッチングさせる。PSの値は両者で近似している。先の右心カテーテルにおいては0.03 以内の範囲でPSのマッチングが行なわれた。

```

. ttest ps, by(allcaiser)
Two-sample t test with equal variances

```

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Intervall]	
0	622	.3215199	.0099511	.2481808	.3019779	.3410618
1	622	.3218356	.0099507	.2481685	.3022946	.3413766
combined	1244	.3216777	.0070335	.2480748	.3078788	.3354766
diff		-.0003157	.0140727		-.0279246	.0272932

```

Degrees of freedom: 1242
Ho: mean(0) - mean(1) = diff = 0
Ha: diff < 0      t = -0.0224      P < t = 0.4911
Ha: diff != 0    t = -0.0224      P > |t| = 0.9821
Ha: diff > 0     t = -0.0224      P > t = 0.5089

```

図11. マッチングを行なったあと、帝王切開例622例、非帝王切開例622例となった。両者のpropensity score の間に有意差はない。

```
. tab chudoku allcaiser, chi
```

chudoku	allcaiser		Total
	0	1	
0	3,762	862	4,624
1	108	151	259
Total	3,870	1,013	4,883

Pearson chi2(1) = 234.6247 Pr = 0.000

↓

```
. tab chudoku allcaiser, chi
```

chudoku	allcaiser		Total
	0	1	
0	540	551	1,091
1	82	71	153
Total	622	622	1,244

Pearson chi2(1) = 0.9018 Pr = 0.342

図12 . Propensity score でマッチングする前は、妊娠中毒症の頻度が明らかに帝王切開例に多い(上)。しかし、マッチングしたあとは、均衡がとれ帝王切開施行例、非施行例の間で妊娠中毒症の併発に差を認めなかった。

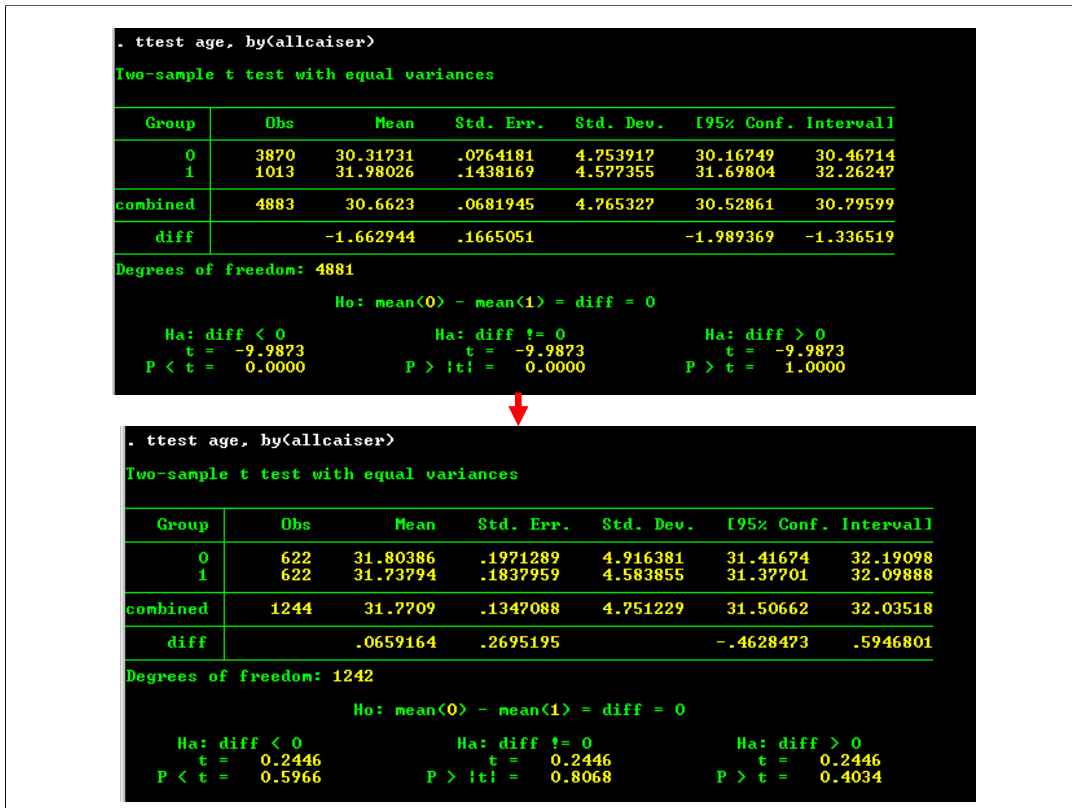


図13. マッチング前は、明らかに妊婦の年齢は妊娠中毒例で高い(上)。しかし、マッチングした後は近似し、両者間で統計学的有意差を認めない。

### 帝王切開を行うケースでは仮死が多い

```
. logistic asph allcaiser
```

Logistic regression		Number of obs	=	4883
Log likelihood = -744.98393		LR chi2(1)	=	25.32
		Prob > chi2	=	0.0000
		Pseudo R2	=	0.0167

asph	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
allcaiser	2.321958	.3723432	5.25	0.000	1.695735 3.179442



### 帝王切開は仮死と無関係

```
. logistic asph allcaiser
```

Logistic regression		Number of obs	=	1244
Log likelihood = -280.12306		LR chi2(1)	=	0.92
		Prob > chi2	=	0.3371
		Pseudo R2	=	0.0016

asph	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
allcaiser	.7939459	.1913911	-0.96	0.338	.4949913 1.273457

図14. 帝王切開と仮死の発生について検討する。マッチング前(上)では、帝王切開により仮死が多くなると判定された。しかし、propensity score でマッチングしたあとは、仮死と帝王切開の間に有意な関係をみいだすことはできなくなっている。

帝王切開と仮死の関係: 19 の変数を用いた多変量解析

```
. logistic asph age bunkai shu sohaku yokon yosui hinmyaku hasui cpd zenchi iugr male bijyaku
> n toui yuhatsu allcaiser nojomyaku single
```

```
Logistic regression          Number of obs   =      4879
                             LR chi2(19)          =      359.38
                             Prob > chi2          =      0.0000
Log likelihood = -574.51539   Pseudo R2       =      0.2383
```

asph	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
age	.9739652	.0179591	-1.43	0.153	.9393945 1.009808
bunkai	1.299762	.1267509	2.69	0.007	1.073633 1.573519
shu	.7327614	.020632	-11.04	0.000	.693419 .7743359
sohaku	4.057638	1.542855	3.68	0.000	1.925825 8.549283
yokon	1.921099	.4061787	3.09	0.002	1.269345 2.907499
yosui	2.080574	.6744884	2.26	0.024	1.102141 3.927615
hinmyaku	13.9131	8.737219	4.19	0.000	4.063327 47.63935
hasui	.9286628	.1836339	-0.37	0.708	.6302914 1.368279
cpd	1.498064	.8140997	0.74	0.457	.5163668 4.34613
zenchi	1.315671	.9297078	0.39	0.698	.3293466 5.255834
iugr	.5217046	.2203986	-1.54	0.124	.2279436 1.194049
male	1.228373	.2127745	1.19	0.235	.8747585 1.724934
bijyaku	1.742928	.4638979	2.09	0.037	1.034483 2.936536
kaisen	1.084911	.4632686	0.19	0.849	.4698099 2.505338
toui	.4088803	.1045748	-3.50	0.000	.2476816 .674992
yuhatsu	1.377068	.3518782	1.25	0.211	.8345456 2.272272
allcaiser	1.552821	.3701106	1.85	0.065	.9732848 2.477439
nojomyaku	.2113248	.0414635	-7.92	0.000	.1438594 .3104293
single	.8296841	.2156982	-0.72	0.473	.4984488 1.381036

図15. 今度は仮死の発生を通常の多ロジスティックモデルで解析してみた。Propensity score を用いた場合同様、帝王切開は有意差を示していない。

帝王切開を行うケースではRDSが多い

```
. logistic rds allcaiser
```

Logistic regression		Number of obs	=	4883
Log likelihood = -225.06141		LR chi2(1)	=	70.61
		Prob > chi2	=	0.0000
		Pseudo R2	=	0.1356

rds	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
allcaiser	14.22313	5.109656	7.39	0.000	7.033984 28.76002



帝王切開はRDSと無関係

```
. logistic rds allcaiser
```

Logistic regression		Number of obs	=	1244
Log likelihood = -89.105699		LR chi2(1)	=	1.51
		Prob > chi2	=	0.2187
		Pseudo R2	=	0.0084

rds	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
allcaiser	1.848336	.9439835	1.20	0.229	.6792892 5.029296

図16. 今度は結果を仮死ではなく、ひり発生頻度の低い呼吸窮迫症候群(RDS)にした。マッチングを行なわないと帝王切開はRDSの発生リスクを上げる。しかし、Propensity score でマッチングしたあとは、この有意差は失われる。



帝王切開とRDSの関係: 13の変数を用いた多変量解析

Logistic regression		Number of obs = 4607			
Log likelihood = -102.38562		LR chi2(13) = 310.58			
		Prob > chi2 = 0.0000			
		Pseudo R2 = 0.6027			
rds	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Intervall]
age	.9612493	.0402549	-0.94	0.345	.8855023 1.043476
bunkai	.8024272	.1953498	-0.90	0.366	.4979455 1.293092
shu	.4946907	.0335249	-10.39	0.000	.43316 .5649618
sohaku	3.494015	2.79109	1.57	0.117	.7300755 16.72175
yokon	.4585187	.3374393	-1.06	0.289	.1083751 1.939924
yosui	.3171869	.2218399	-1.64	0.101	.0805353 1.249234
kyomyaku	1.247761	.2901072	0.95	0.341	.7910882 1.968057
twin	.9094038	.2738673	-0.32	0.753	.5039844 1.640954
hasui	1.117171	.5194949	0.24	0.812	.4490586 2.779306
zenchi	6.882859	5.646061	2.35	0.019	1.378866 34.35705
iugr	.6662115	.5324502	-0.51	0.611	.1390958 3.190878
toui	.5989786	.2850389	-1.08	0.281	.2356925 1.522218
allcaiser	3.909084	2.025899	2.63	0.009	1.415586 10.79478

図17. 今回は多ロジスティック解析を行なった。前回の仮死の場合とは異なり、一番下の変数である帝王切開は有意性を示している。

以上のように、結果発生が変数 x 7 より大きかった仮死の場合には、多ロジスティック解析を用いた場合にも、propensity score を用いた場合にも帝王切開が仮死に影響しないという点で一致していた。

しかし、呼吸窮迫症候群という少ない結果を選んだとき、その数は変数 x 7 を超え、Am J Epi の指摘した通りとなった。すなわち、「結果発生に対して選択する変数の数が極端に少ない場合、propensity score を用いるべきである」点である。