

# Confounding Potential Plot

Drew Griffin Levy, PhD  
Senior Medical Science Director  
Genentech, Inc.  
Evidence Science & Innovation | US Medical Affairs  
[levy.drew@gene.com](mailto:levy.drew@gene.com)

December 13, 2016

**Very Low Birth Weight Infants dataset<sup>1</sup>**  
**32 Variables      644 Observations**

---

**birth**  
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95  
644 0 537 1 84.75 1.838 82.05 82.45 83.51 84.90 86.07 86.90 87.19

lowest : 81.511 81.514 81.552 81.558 81.593, highest: 87.455 87.466 87.469 87.472 87.483

---

**exit**  
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95  
636 8 535 1 84.85 1.922 82.19 82.56 83.59 84.96 86.17 87.01 87.34

lowest : 68.526 81.046 81.169 81.539 81.552, highest: 87.718 87.721 87.800 87.838 96.871

Value 68.5 81.0 81.5 82.0 82.5 83.0 83.5 84.0 84.5 85.0 85.5 86.0 86.5 87.0  
Frequency 1 2 9 25 42 46 60 42 62 72 70 57 58 52  
Proportion 0.002 0.003 0.014 0.039 0.066 0.072 0.094 0.066 0.097 0.113 0.110 0.090 0.091 0.082

Value 87.5 88.0 97.0  
Frequency 35 2 1  
Proportion 0.055 0.003 0.002

---

**hospstay**  
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95  
636 8 152 1 40.21 83.14 1 3 16 37 62 98 129

lowest : -6574 -295 -293 -291 -288, highest: 276 300 442 797 3668  
-6600 (1, 0.002), -300 (4, 0.006), -200 (2, 0.003), 0 (412, 0.648), 100 (193, 0.303), 200 (18, 0.028), 300 (3, 0.005), 400 (1, 0.002), 800 (1, 0.002), 3700 (1, 0.002)

---

**lowph**  
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95  
604 40 73 0.999 7.202 0.157 6.94 7.01 7.13 7.21 7.31 7.36 7.39

lowest : 6.529999 6.699997 6.719997 6.739998 6.759998  
highest: 7.480000 7.489998 7.500000 7.519997 7.549999

---

**pltct : platelet count**  
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95  
597 47 266 1 201.5 90.72 69.6 97.2 143.0 202.0 252.0 304.8 337.0

lowest : 16 24 28 34 38, highest: 399 416 418 462 571

---

**race**  
n missing distinct  
639 5 4

white (255, 0.399), black (364, 0.570), native American (16, 0.025), oriental (4, 0.006)

---

**bwt : birth weight [g]**  
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95  
644 0 137 1 1104 296.1 641.5 741.5 910.0 1135.0 1320.0 1440.0 1480.0

lowest : 400 430 460 480 500, highest: 1475 1480 1490 1500 1580

---

**gest : gestational age [weeks]**  
n missing distinct Info Mean Gmd .05 .10 .25 .50 .75 .90 .95  
643 1 24 0.987 28.98 2.767 25 26 27 29 31 32 33

lowest : 22.0 23.0 24.0 25.0 25.5, highest: 34.0 35.0 36.0 38.0 40.0

---

**inout**  
n missing distinct  
644 0 2

born at Duke (528, 0.82), transported (116, 0.18)

---

**twn**  
n missing distinct  
644 0 2

FALSE (509, 0.79), TRUE (135, 0.21)

<sup>1</sup>See O'Shea M, Savitz DA, Hage ML, Feinstein KA: Prenatal events and the risk of subependymal / intraventricular haemorrhage in very low birth weight neonates. Paediatric and Perinatal Epidemiology 1992;6:352-362. See also, <http://biostat.mc.vanderbilt.edu/wiki/pub/Main/DataSets/vlbw.html>.

---

**lol**

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
288	356	40	0.958	8.497	12.23	0.00	0.00	0.00	4.00	9.00	20.00	26.65

lowest : 0 1 2 3 4, highest: 72 102 104 174 192

---

**magsulf**

n	missing	distinct	Info	Sum	Mean	Gmd
420	224	2	0.347	56	0.1333	0.2317

---

**meth**

n	missing	distinct	Info	Sum	Mean	Gmd
563	81	2	0.738	246	0.4369	0.4929

---

**toc:** tocolysis - mother treated with  $\beta$ -adrenergic drug

n	missing	distinct	Info	Sum	Mean	Gmd
563	81	2	0.524	127	0.2256	0.35

---

**delivery**

n	missing	distinct
644	0	2

abdominal (312, 0.484), vaginal (332, 0.516)

---

**apgl**

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
631	13	10	0.985	4.903	3.01	1	1	2	5	7	8	9

lowest : 0 1 2 3 4, highest: 5 6 7 8 9

0 (5, 0.008), 1 (91, 0.144), 2 (68, 0.108), 3 (57, 0.090), 4 (45, 0.071), 5 (68, 0.108), 6 (79, 0.125), 7 (79, 0.125), 8 (102, 0.162), 9 (37, 0.059)

---

**vent**

n	missing	distinct	Info	Sum	Mean	Gmd
635	9	2	0.731	368	0.5795	0.4881

---

**pneumo**

n	missing	distinct	Info	Sum	Mean	Gmd
640	4	2	0.477	127	0.1984	0.3186

---

**pda**

n	missing	distinct	Info	Sum	Mean	Gmd
637	7	2	0.498	134	0.2104	0.3327

---

**cld**

n	missing	distinct	Info	Sum	Mean	Gmd
598	46	2	0.588	160	0.2676	0.3926

---

**pvh**

n	missing	distinct
513	131	3

absent (349, 0.68), possible (41, 0.08), definite (123, 0.24)

---

**ivh**

n	missing	distinct
514	130	3

absent (430, 0.837), possible (10, 0.019), definite (74, 0.144)

---

**ipe**

n	missing	distinct
514	130	3

absent (461, 0.897), possible (17, 0.033), definite (36, 0.070)

---

**year**

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
644	0	537	1	84.75	1.838	82.06	82.45	83.52	84.90	86.07	86.90	87.19

lowest : 81.51196 81.51471 81.55304 81.55847 81.59406  
highest: 87.45721 87.46814 87.47089 87.47363 87.48456

---

**sex**

n	missing	distinct
643	1	2

female (317, 0.493), male (326, 0.507)

---

**dead**

n	missing	distinct	Info	Sum	Mean	Gmd
644	0	2	0.466	124	0.1925	0.3114

---

**cesarean**

n	missing	distinct
644	0	2

FALSE (332, 0.516), TRUE (312, 0.484)

---

**definite.ivh**

n	missing	distinct
514	130	2

FALSE (440, 0.856), TRUE (74, 0.144)

---

**white**

n	missing	distinct
639	5	2

FALSE (384, 0.601), TRUE (255, 0.399)

---

**male**

n	missing	distinct
643	1	2

FALSE (317, 0.493), TRUE (326, 0.507)

---

**txported**

n	missing	distinct
644	0	2

FALSE (528, 0.82), TRUE (116, 0.18)

---

**yyyy**

n	missing	distinct
644	0	7

lowest : 1981 1982 1983 1984 1985, highest: 1983 1984 1985 1986 1987

1981 (29, 0.045), 1982 (83, 0.129), 1983 (107, 0.166), 1984 (122, 0.189), 1985 (130, 0.202),  
1986 (119, 0.185), 1987 (54, 0.084)

---

Table 1: Descriptive Statistics by delivery

	N	abdominal	vaginal
		$N = 312$	$N = 332$
gestational age <small>weeks</small>	643	29	29
birth weight <small>g</small>	644	1172	1100
twn	644	28%	14%
white	639	46%	34%
inout : transported	644	11%	24%
male	643	50%	52%
yyyy : 1981	644	4%	5%
1982		12%	13%
1983		18%	15%
1984		16%	22%
1985		21%	20%
1986		19%	18%
1987		10%	7%

$a$   $b$   $c$  represent the lower quartile  $a$ , the median  $b$ , and the upper quartile  $c$  for continuous variables.  $N$  is the number of non-missing values.

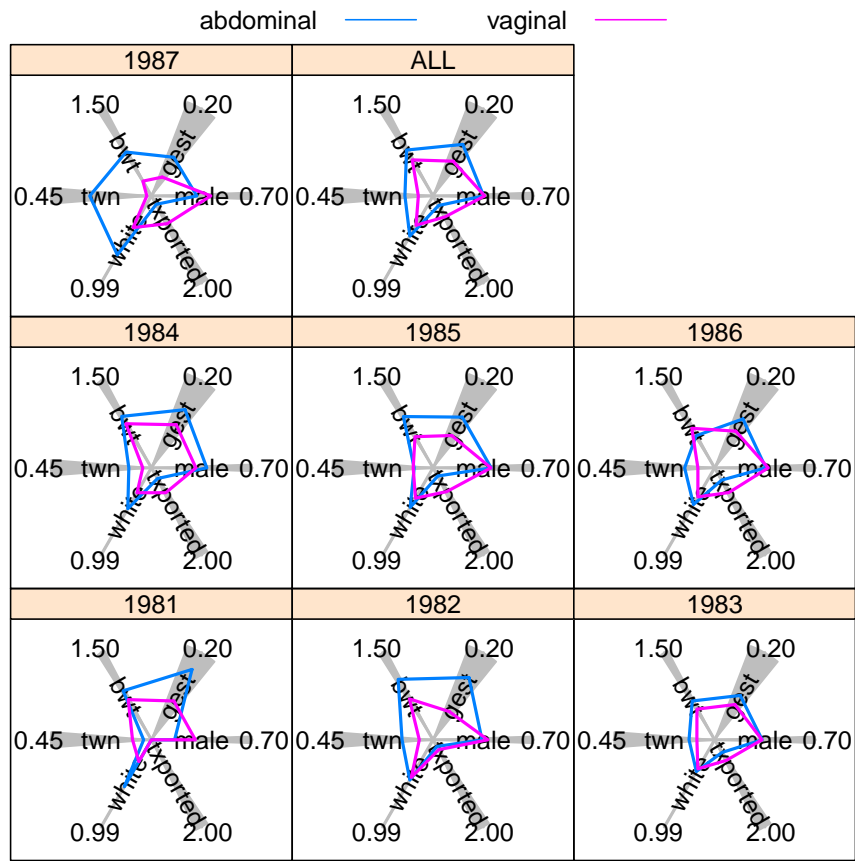


Figure 1: **Trellised radar plot.** The ALL panel can be seen to correspond to the ‘standard Table 1’ shown in Table 1 above. The trellis can be seen therefore to produce a disaggregated ‘Table 1’, in this case exploring the possibility of a secular trend in potential confounding.