A short walk in human reproduction

N Mascie- Taylor [ab] & JC Thalabard [acd]
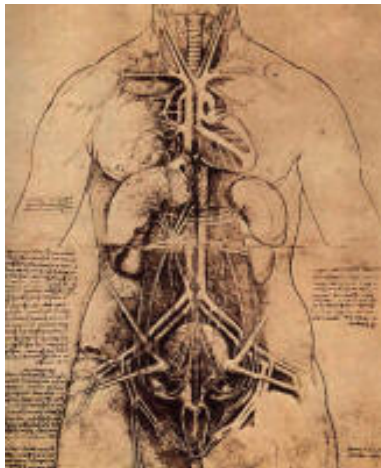
[a]Churchill College,
[b]Dept Biological Anthropology, Cambridge Univ.,
[c]MAP5 UMR CNRS 8145, University Paris Descartes, SPC, [d]Hôpital Cochin
PR1, APHP

20/07/2014

The Reproductive Function : a very special physiologic function

Leonardo da Vinci (1452- 1519) and the utero- mammary duct : when you see what you expect !

# The Reproductive Function : a very special physiologic function

- ▶ not essential for the individual's survival...but essential for the group survival
- ▶ Changes in reproductive behavior concern societies and their evolution
    - ▶ The one- child policy and the gender selection
    - ▶ The delayed child bearing age in westernised countries and the demographic changes
    - ▶ A "cultural weapon"

How to optimize mating both quantitatively and qualitatively ?

- A very old question, which attracted many famous scientists from all fields
  - W Harvey 1651 : *Exercitationes de Generatione Animalium* edited by George Ent
  - GLL Buffon 1749 : Histoire naturelle. Life tables (Le Bras, 2000)
  - G Mendel (1822- 1884)
- The Reproductive Function and mathematical modeling
  - Usually driven by agricultural issues for improving efficiency (cultures, herding)
  - The sex ratio issue and the arguments between *J Arbuthnot, Bernoulli DeMoivre*
  - Callipedia, Orthopedia, Eugenism...
  - The study of complex psychological traits, intelligence (IQ)
  - The problem of the extinction of the family names and the 19th century context : *the Galton- Watson process*
- Any statistical modeling in reproduction needs to define both the correct **statistical unit** and an adapted **time scale**

The Reproductive Function : a quick reminder (XY excluded !)

- The fallacy of menstrual bleeding
- The hidden endocrine system
- A complex system regulated by both internal and external stimuli
    - A subtle dynamic interplay between 3 entities
    - Each one has its own timescale and tempo
- Two important landmarks : menarche and menopause
- A dynamical system with reversible and irreversible states
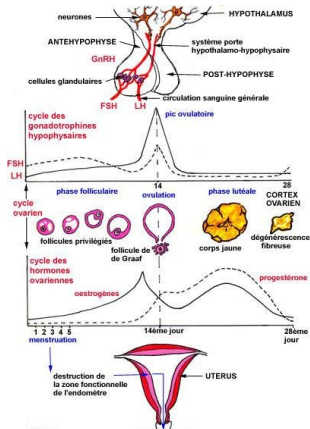
The neuro- endocrine control of the pituitary- gonadal axis

- ▶ G Harris and the beginning of neuroendocrinology (Kreier and Swaab, 2010) : the retino- ejaculatory reflex in ducks
- ▶ Multiple signals, both endogenous and exogenous, mostly inhibitory
  - ▶ Light/ Stressors/ Olfactory stimuli
  - ▶ Steroid sexual hormones, both endogenous and exogenous (contraception, *endocrine disruptors* -cf. Obseff)
  - ▶ Insulin, Leptin in relation to nutritional and metabolic stimuli
  - ▶ Immune System (IL) and infectious/ chronic diseases
  - ▶ *Breastfeeding* R Short and the WHO study(P Howie et al), cf Bangladesh tea plantation study NMT & LR, (LeStrat and Thalabard, 2001)

Example 1 : The Obseff study. Context and aims

- Context and aims
    - Participants : A Bohet, J Bouyer, B Ducot, N Keiding, **L Rosetta**, **R Slama (PI)**, JCT
    - Current issue : is there a decline in fertility in Europe in relation to environment ? $\Rightarrow$ **Obseff main objective** : assessing the actual fertility and its trend in relation to environment in French couples **from the general population** wishing to conceive (Slama et al., 2012)
- Ancillary study : what is the female ovulatory status in those couples ?
    - Its evidence usually based in clinical practice on a set of complementary tools
        - Menstrual Diary, Body Temperature Chart, Follicular US Monitoring
        - $\pm$ Repeated Collections of either Blood or Salivary or Urinary Individual Samples
    - QS1 : Is urinary repeated collection an acceptable tool in **Field Studies ?**
    - QS2 : Can we reduce the urinary frequency collection without compromising ovulation assessment ?
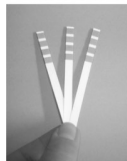
Example 1 : The Obseff study : a complex survey

- ▶ 64,805 homes contacted by random digit dialling
- ▶ 15,811 women identified 18-44 yrs $\propto$ eligibility questionnaire
- ▶ 997 with regular intercourse, without any contraception, no delivery within previous 6 mo
    - ▶ 685 accepted hormonal assessment $\rightarrow$ 250 complied with protocol including ethical agreement + booklet filled in
    - ▶ 41 only filled in the booklet
    - ▶ 209 accepted urinary sampling. Among them, 49 women **randomly** sampled for every day sampling, whereas 160 collected every other day.
    - ▶ 15 pregnancies reported

Example 1 : The Obseff study. Urinary collection and E1G/PdG determinations

- Urine collection
  - Set of calibrated filter papers for a full menstrual cycle
  - Self- collection **every- or every other- day**
  - Each measure = 3 pH strips $\times$ 4 squares (1 square $\simeq$ 5 μ l)
  - Strips exposed to 1st urine flow after wake up, dried at room θ, stored into a plastic bag.
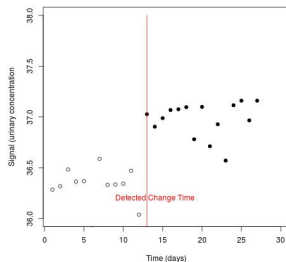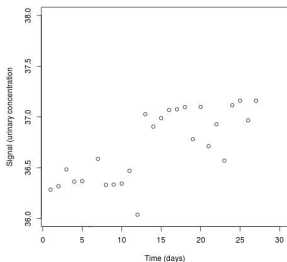 - All plastic bags for a cycle sent by regular mail to central lab

- Urinary assay
  - Assay in duplicate PdG, EiG, Creat
 - Procedure $\propto$ WHO procedure using WHO reagents

Example 1 : the Obseff study : Ovulation detection

▶ The mathematical problem relates to the detection of a signal change in a time series

  ▶ The distinction on- line $\neq$ **off- line** detection but a situation of **sparse** series
▶ Several methods reported (Santoro et al., 2003, Thalabard et al., 2011)
▶ Validation : 3 independent experts analysed blindly 96 anonymous profiles in random order (2×48)

Example 1 : Obseff study. Some PdG Profiles
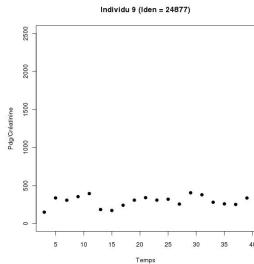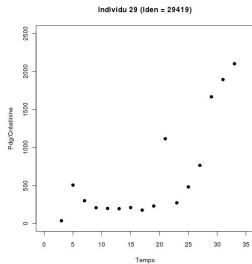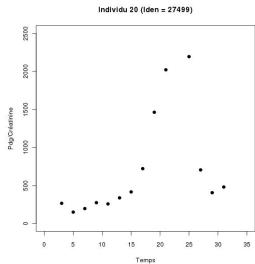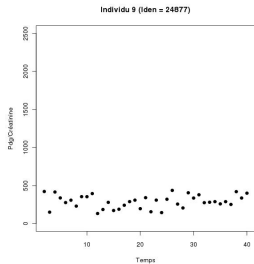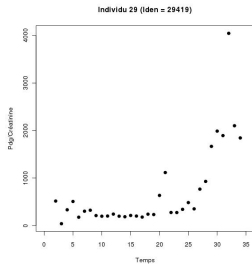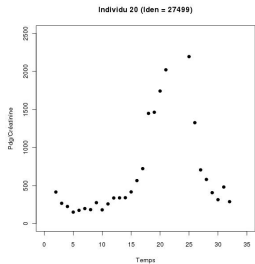
Example 1 : Obseff study. Did the experts agree ?

- ▶ Ovulation Discordances

Frequency	Exp 1 vs 2	Exp 1 vs 3	Exp 2 vs 3
1	2	1	1
2	3	1	2

- ▶ Day of Ovulation

Frequency	ICC	95%CI		Mean 2-2 CC	Robinson	Finn
1/1	0.90	0.84	0.95	0.91	0.94	0.95
2/1	0.94	0.89	0.96	0.91	0.93	0.95
2/2	0.84	0.76	0.90	0.86	0.90	0.96

Example 1 : Obseff study. How the automatic detectors agreed with the experts ?

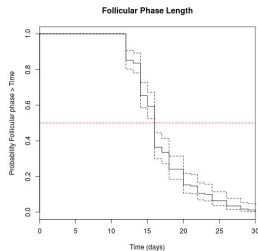| Frequency	Method	Se	Sp	True Ovulating	Non Ovulating
1	Brown	1.00	0.50	44	4
1	Kassam	0.93	0.75	44	4
1	Waller1	0.93	0.75	44	4
1	Black95	1.00	0.25	44	4
1	Black99	1.00	0.00	44	4
1	NLR	1.00	0.25	44	4
1	MBP	1.00	0.25	44	4
1	PWR	1.00	0.25	44	4
1	Cusum	1.00	0.25	44	4
1	Ewma	0.84	0.75	44	4
2	Brown	0.93	0.6	43	5
2	Kassam	1.00	0.8	43	5
2	Waller1	0.98	0.8	43	5
2	Black95	0.98	0.4	43	5
2	Black99	0.98	0.4	43	5
2	NLR	1.00	0.2	43	5
2	MBP	1.00	0.4	43	5
2	PWR	1.00	0.2	43	5
2	Cusum	1.00	0.4	43	5
2	Ewma	0.35	0.8	43	5

Example 1 : Obseff study. Cycle characteristics in couples from the general population
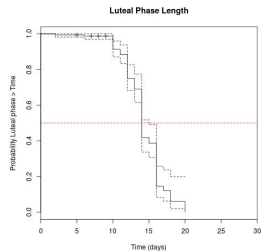
The Kaplan Meier representation for **right**censored data
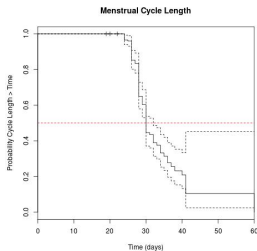
Cycle Length   Follicular Phase   Luteal Phase

Discussion and conclusions

- ► Urinary self collection appears feasible
    - ► In field situation the choice of urine collection on pH strips appears well adapted & well accepted
    - ► Few missing points as soon as the woman accepts starting
    - ► No need of storage in the fridge = well accepted by family !
    - ► The whole set not cumbersone, sent by post to the lab, reliable, not costly
    - ► Enzyme immuno assay not expensive and friendly for the environment
- ► Main Results
    - ► 3 experts gave consistant results with no doubtful situations
    - ► The simple creat- corrected PdG threshold Kassam's one gave the "best" results
    - ► Use of every other day versus every day collection : the range of variation remained small and acceptable for **off line detection** in population study

Example 2 : Is mating for height associated with fertility ?

current work with Monika Krzyzanowska and NMT

- ▶ How to select his or her partner remains a burning issue !
    - ▶ From the oracle's advice and regional recipes to the current illusion of scientifically mastering an event, which is indeed largely influenced by the socio- economic context, the cultural environment and chance.
    - ▶ What has really changed ? (Allen, 1997)
- ▶ QS : are basic biometric parameters still important for fertility ?

Example 2 : Revisiting an old question. An intriguing graph

The UK National Child Development Study (NCDS) data set

- Children born March, 3rd-9th 1958. Regular FU : 1965, 1969, 1991, 2000, 2004, 2008, 2013.

- 1974 FU : 6535 index children, with information on age, weight and height, education and social class in both parents.
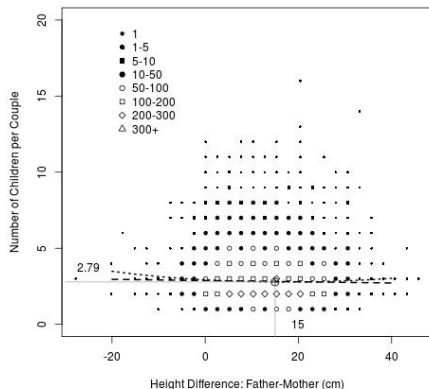
Example 2 : The choice of the regression model

- ▶ Here, Y, the number of children is integer- valued $\Rightarrow$ a **counting process**

- ▶ In addition, only couples with at least one child are recorded : Y is **left- truncated**

- ▶ Linear regression analysis tries to explain a dependent variable Y with a set of independent regressors X_1, X_2 *as parcimonious as possible*

- ▶ Potential regressors are either continuous like height, weight or categorical with a natural order like education- or social- scores or without like regions

- ▶ The culturally- linked regressor variables are expected to be somehow highly correlated : "qui se ressemble s'assemble !"
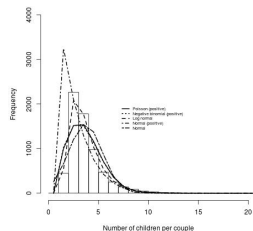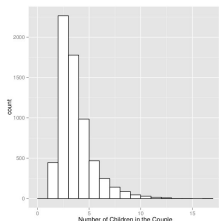
Example 2 : adjusting on covariates

Covariate		Estimate	p-value
Region	Scotland (Reference)	1.66	**1.73e-11**
Δ	North	-0.06	8.34e-02
	Yorks+Humber	-0.12	**1.17e-03**
	East Midlands	-0.08	3.36e-02
	East Anglia+South East	-0.13	**5.12e-07**
	South West	-0.06	7.02e-02
	West Midlands	-0.08	**7.27e-03**
	North West	-0.03	3.47e-01
	Wales	-0.05	1.740e-01
Age	mean	0.02	**3.03e-57**
	difference	-0.00	1.35e-01
Social class	mean	0.14	**2.31e-29**
	difference	0.02	**3.48e-04**
Education	mean	0.04	**4.64e-04**
	difference	0.01	2.77e-01
Height	mean	-0.01	**1.48e-11**
	difference	**0.001**	**1.76e-01**

The height difference is no more significant, but the social difference seems
important

Example 2 : the unexpected effect of the difference in social class
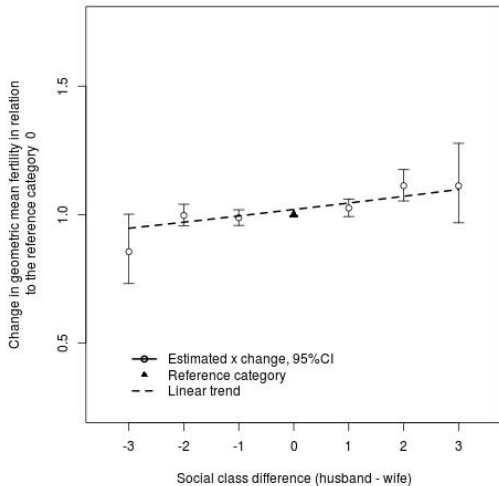
Example 3 : Determinants of age at menarche in UK

current work with Monika Krzyzanowska and NMT

- ▶ Age at menarche is an important natural event in a woman's life occuring on a continuous age time- scale
- ▶ But its precise value can be altered by **right censoring**,recall bias and rounding bias leading to **interval censored** data

⇒ A very frequent situations in clinical and field studies

Real values   Right censoring   Interval- censoring
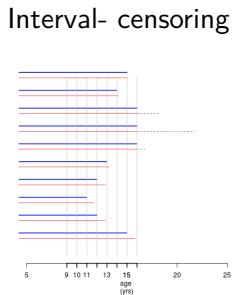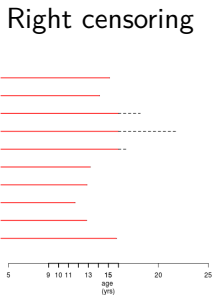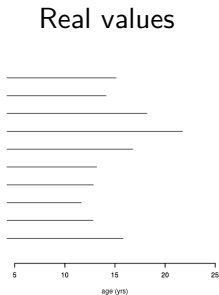
Example 3 : Determinants of age at menarche in UK

- ▶ Classical tool for right censored survival data : Cox semi-parametrice proportional hazard regression model
  - ▶ Continuous time scale, few, if any, tied values
  - ▶ Asymptotic results now connected to the counting process theory
- ▶ Alternative models
  - ▶ Parametric models, Accelerated Failure Time (AFT) model
  - ▶ Poisson regression for counting data
- ▶ Direct interval censored survival data analysis
  - ▶ Peto (1973) pointed out the issue precisely in relation to menarche data !
  - ▶ Recent algorithmic developments are now accessible (Gomez et al., 2009)

Example 3 : Some very cautious preliminary results

Origin	Variable		Estimate	p-value
Daughter	**Weight (kg) at 7 yrs**		0.067	$0.0e+00$
	Height (cm) at 7 yrs		0.03	$0.0e+00$
	Breast Fed	1/2/3		NS
Mother	**Age at menarche**		-0.15	$0.0e+00$
	Educ cat	2/3/4		NS
	Smoking during Pregnancy	Yes		NS
Father	Educ. cat	2/3/4		NS
	Mobility 58-65 7	0/1		NS
	Financial Diff.	Yes	0.20	$2.8e-35$
Family	Social cat	1/2/3/		NS
	Birth order	2	-0.04	$1.6e-01$
		3	-0.09	$3.0e-04$
	Number of Children		-0.08	$2.0e-67$
Environment	Crowding	$1-1.5$	-0.1288	$6.1e-06$
		$1.5-2$	-0.2324	$2.7e-07$
		2+	-0.2843	$3.1e-04$
	Bedroom share	2	-0.092	$1.3e-04$
		3	-0.23	$2.0e-09$
	Amenities at 7			NS
	Tenure at 7	2	-0.10	$1.1e-05$
		3	0.03	$5.0e-01$
		4	0.121	$6.1e-02$
	Free School Meals	Yes	0.29	$1.2e-76$
	Accommodation	1/2/3		NS
Region	North		-0.051	$3.1e-01$
	Yorks.Humber		0.048	$3.6e-01$
	East.Midlands		0.02	$7.0e-01$
	East.Anglia.SouthWest		0.18	$2.1e-11$
	SouthWest		0.11	$5.9e-02$
	West.Midlands		0.17	$4.3e-04$
	NorthWest		0.01	$8.3e-01$
	Wales		0.22	$4.1e-04$

Politics, environment and reproduction : a short story before conclusion...

- ▶ Do you recognize this plant ?

- Heracleum Sphondylyum : (common) hogweed/ berce (commune)/ Bearfoot...
- Old popular knowledge about its side effects but also its health benefits

- Very abundant in Scotland, where the natality is high... $\Rightarrow$ Queen Victoria is reported to have mandated her chemists to design one of the first herbicide against hogweed ?

$\Rightarrow$ Of course, correlation is not causation, but it frequently drives our decisions... and, if the story is correct, it visibly failed here !

Some concluding remarks

This 6- month overseas fellowship at Churchill was a unique opportunity

- ▶ to find more time to work on continuously postponed scientific tasks
- ▶ to enjoy the Churchill College life and the Cambridge fantastic environment
- ▶ to meet extraordinary people from so different fields in such a rather short period, something we are not very successful at SPC !
- ▶ to discover unexpected topics of interest !
- ▶ to foster scientific collaborations and envisage joint meetings (MD/ PhD program)
- ▶ to enjoy teaching practical use of R- based biostatistics for field studies to a very diversified audience

I sincerely thank all the people both at the French Embassy, at Churchill and at Nick's lab who made this stay possible and so enjoyable !

References I

G. E. Allen. The social and economic origins of genetic determinism : a case history of the american eugenics movement, 1900-1940 and its lessons for today. *Genetica*, 99 :77–88, 1997.

G. Gomez, M. L. Calle, R. Oller, and K. Langohr. Tutorial on methods for interval-censored data and their implementation in r. *Statistical Modelling*, 9 :259–297, 2009.

F. Kreier and D. F. Swaab. Chapter 23 : history of neuroendocrinology "the spring of primitive existence". *Handb Clin Neurol*, 95 :335–360, 2010.

H. Le Bras. *Naissance de la mortalité. L'origine politique de la statistique et de la démographie*. Seuil/ Gallimard, 2000.

Y. LeStrat and J. Thalabard. Analysis of postpartum lactational amenorrhoea in relation to breast-feeding : some methodological and practical aspects. *J Biosoc Sci*, 33(4) :529–549, Oct 2001.

References II

R. J. Peto. Experimental survival curves for interval-censored data. *Journal of the Royal Statistical Society, Series C*, 22 :86–91, 1973.

N. Santoro, S. L. Crawford, J. E. Allsworth, E. B. Gold, G. A. Greendale, S. Korenman, B. L. Lasley, D. McConnell, P. McGaffigan, R. Midgely, M. Schocken, M. Sowers, and G. Weiss. Assessing menstrual cycles with urinary hormone assays. *Am J Physiol Endocrinol Metab*, 284 :E521–E530, 2003.

R. Slama, O. Hansen, B. Ducot, A. Bohet, D. Sorensen, E. M. Giogis Allemand, L, L. Rosetta, J. Thalabard, and N. Keiding. Estimation of the frequency of involuntary infertility on a nation-wide basis. *Hum Reprod*, 27 :1489–1498, 2012.

J. Thalabard, L. Joubin, L. Rosetta, and C. Mascie-Taylor. *Analysing the characteristics of the menstrual cycle in field situations in humans : some methodological aspects*. Cambridge University Press, 2011.