Protection of health information in Italy: a step too far?

From MARINA CUTTINI,[1]* CRISTIANO MARINI,[1] SILVIA BRUZZONE,[2] SABRINA PRATI[2] and RODOLFO SARACCI[3]

A recent appeal by a group of Italian obstetricians and neonatologists, advocating full resuscitation of extremely preterm infants independently from parental opinion, raised a debate on the rationale and consequences of such proposal.[1]

Whether or not the appeal will modify practices, there is no doubt that careful assessment of outcome for these very special infants is called for. However, this is currently impossible at national level in Italy. Following a change in legislation,[2] the time-honoured system of births monitoring by the Italian National Institute of Statistics (ISTAT) was dismantled in 1998 and later rebuilt entrusting it to the Ministry of Health, while ISTAT remains in charge of deaths registry. Both are public institutions; yet for privacy protection the transfer of birth certificates from the Ministry to ISTAT is only permitted after deletion of personal identifiers. Thus, the individual matching of birth certificates, containing crucial information such as birthweight, gestational age and vitality, to the corresponding infant death data (if any) becomes more difficult.

Results of an attempt of statistical record linkage[3] performed on the 2003 birth cohort are shown in Table 1. Variables used as keys to record linkage were infant's gender, plurality, date and place of delivery and maternal date of birth. Overall, only 598 of the 1539 deaths (38.9%) could be successfully linked, and proportions decreased from 56.6% in the North to 32% in the Centre and South, to only 15% in the Islands. Both missing birth certificates and records incompleteness on linkage variables contributed to these results. The lower proportion of valid death records determined linking difficulties in Central Italy, while missing birth certificates were the main issue in the South and especially Islands. At the light of the North-to-South trend of neonatal and infant mortality traditionally reported in Italy, lack of information from the Southern regions is particularly troubling.

As stated by the Europeristat project, neonatal and infant mortality stratified by birthweight and gestational age are 'core' indicators to be recorded by all European Union countries to assess the quality of perinatal care and monitor the effects of policy changes.[4] Voluntary collection of data by Neonatal Intensive Care Units, as developing today in Italy and other countries for benchmarking purposes,[5] is a useful but inadequate substitute, being based on the selected subgroup of neonates surviving to admission to tertiary Centres.

Solutions are urgently needed to reconcile privacy protection with timely population-based monitoring of neonatal and infant outcomes stratified by birthweight and gestational age.

Conflict of interest: None declared.

Table 1 Results of statistical record linkage performed on 2003 birth cohort

Region	Livebirths		Neonatal Deaths			Record linkage		
	Number of completed birth certificates	Percentage of completed birth certificates on number of livebirths[a]	Number of neonatal deaths	Number and percentage of valid death records[b]		Number of linked death records	Percentage of linked death records on valid records[b]	Percentage of linked death records on total deaths
	No.	(%)	No.	No.	(%)	No.	(%)	(%)
Northern Italy	222 978	95.4	581	388	66.8	329	84.8	56.6
Central Italy	98 955	100.0	288	125	43.4	93	74.4	32.3
Southern Italy	106 821	75.4	443	275	62.1	142	51.6	32.1
Islands (Sardinia and Sicily)	24 833	37.9	227	145	63.9	34	23.4	15.0
Italy	453 587	84.1	1539	933	60.6	598	64.1	38.9

[a]The number of livebirths (not shown in table) was derived from civil registration.
[b]Valid death records are those where all the linkage variables were validly filled in.

[1] Unit of Epidemiology, Bambino Gesù Children's Hospital, Rome, Italy.
[2] Italian National Institute of Statistics, Rome, Italy.
[3] IFC - National Research Council, Pisa, Italy.
* Corresponding author. Unit of Epidemiology, Bambino Gesù Children's Hospital, Rome, Italy. E-mail: cuttini@opbg.net

References

[1] De Bac M. Se il feto prematuro è vivo va rianimato. [When the preterm fetus is born alive, resuscitation is mandatory] Corriere della Sera. Available at: http://archiviostorico.corriere.it/2008/febbraio/03/feto_vivo_rianimato_scontro_co_9_080203181.shtml (Accessed February 3, 2008).

[2] Law 127/97, *Gazzetta Ufficiale* 113, 17th May 1997 and Executive Regulations 403/98, *Gazzetta Ufficiale* 275, 24th November 1998.

[3] Fellegi IP, Sunter AB. A theory for record linkage. *J Am Stat Assoc* 1969;**64:**1183–210.

[4] Buitendijk S, Zeitlin J, Cuttini M, Langhoff-Roos J, Bottu J. Indicators of fetal and infant health outcomes. *Eur J Obstet Gynecol Reprod Biol* 2003;**111 (Suppl 1):** S66–77.

[5] Valls-i-Soler A, Carnielli V, Claris O *et al*. and Scientific Steering Committee of EuroNeoStat. EuroNeoStat: a European information system on the outcomes of care for very-low-birth-weight infants. *Neonatology* 2008; **93:**7–9.

doi:10.1093/ije/dyn185
Advance Access publication 4 September 2008

Heterogeneous views on heterogeneity

From NIKOLAOS A PATSOPOULOS, EVANGELOS EVANGELOU and JOHN PA IOANNIDIS*

The insightful and stimulating commentary by Julian Higgins[1] on our paper[2] raises several important issues that need to be clarified. First, we need to agree on nomenclature. The heterogeneity literature has been plagued by inconsistent terminology. Terms like 'heterogeneity', 'inconsistency', 'variation', 'diversity', 'between-study variance', 'variability', etc. are used interchangeably. While Higgins prefers the term 'inconsistency' for I^2, in other writings he has used the words 'variability' and 'heterogeneity' in association with this measure.[3] We believe that the term 'heterogeneity' is a nice word with roots going back to ΕΤΕΡΟΓΕΝΗΣ of Aristotle and ΕΤΕΡΟΓΕΝΩΣ of Sextus Empiricus. It can be applied to any of the popular metrics and tests, but then one simply has to specify which metric or test is exactly alluded to. 'Inconsistency' is also a nice, more recent word, but again we need to clarify what it refers to each time.

Higgins worries about 'the *post hoc* hypotheses that need to be thought up to explain why the excluded studies might be outlying or influential'. We were clear cut in our paper that this is indeed not an easy task. We believe that sensitivity analyses, as currently performed, are usually an invitation to *post hoc* data dredging with few or no rules in the game. This reduces their inferential reliability. However, this is a major reason why our proposed algorithms may offer one way to improve this free-lunch situation. There are two components to any sensitivity analysis. The first component is how it is done. The second component is how the results are interpreted. We argue that our method takes away much of the subjectivity in the first component. We do not wish to diminish the uncertainty that arises in the second component, and we wish that all meta-analysts recognize and acknowledge this uncertainty properly.

Higgins questions whether it is sensible to define a 'desired' threshold in terms of I^2 statistic. Although we agree that indeed '(some) heterogeneity is to be expected in (almost any) meta-analysis' and 'any amount of heterogeneity is acceptable, providing both that the predefined eligibility criteria for the meta-analysis are sound and that the data are correct', we believe that using thresholds to describe heterogeneity is an unavoidable consequence of the effort to translate statistical terms into real life. Higgins and colleagues have faced this problem, similarly recommending categorization of values for I^2 and assigning adjectives of low, moderate, and high heterogeneity or inconsistency.[4,5] In our article we have used these values of 50% and 25% for I^2, as traditional thresholds for large and moderate heterogeneity, respectively. This does not negate the need to recognize the major uncertainty in heterogeneity estimates,[6] but provides a standardized approach that can be applied consistently across meta-analyses.

Higgins argues in favour of using τ^2, the estimate of between-study variance, rather than I^2 in our paper, because I^2 depends also on the within-study precisions. Actually I^2 has become popular as a measure primarily due to the groundbreaking work of Higgins.[3,4] I^2 is one of the most commonly reported heterogeneity (or inconsistency) metrics, while the between-study variance τ^2 is rarely reported in the medical literature. I^2 has an intuitive interpretation, and it is comparable across meta-analyses with different numbers of studies or different types of effect metrics, whereas τ^2 is difficult both to understand and compare, according to Higgins' writings.[2] Therefore, we focused on I^2 in our paper. However, the algorithms that we have proposed are not applicable only to I^2. These are general methods that can be used with any kind of metric, e.g. τ^2. If another metric may be useful to apply more widely, we

* Corresponding author. Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina 45110, Greece. E-mail: jioannid@cc.uoi.gr