Review of: "Non-Assisted Hatching Trophectoderm Biopsy Does Not Increase The Risks of Most Adverse Maternal and Neonatal Outcome and May Be More Practical for Busy Clinics: Evidence From China"

Ulrik Schioeler Kesmodel

Potential competing interests: The author(s) declared that no potential competing interests exist

The paper by Li et al. constitutes a well-written and intriguing piece of work within the field of potential risks associated with biopsies from the developping embryo; here trophectoderm biopsies.I particlarly appreciate the size of the dataset and the fairly meticulous methodological description, which is easy to follow.

A few caveats are warranted, though. First, the authors decided to exclude stillbirths and fetuses (or children?) suffering ‘premature death’. The latter term is not defined, so it may include both late miscarriages, late induced abortions, stillbiths, and children born prematurely /preterm but dies shortly after death. Some definition of ‘premature death’ would have been welcomed. Stillbirth and - if relevant - very perterm birth - is associated with adverse pregnancy and birth outcomes, and exclusion may introduce selection bias when inclusion into the study is based on prediction on the future, i.e. livebirth.

Second, mosaic embryos were excluded. It is not obvious how the authors could know with reasonable certainty, which embryos were mosaic, and which were not?

Third, serum-hCG was measured on different gestational days. The authors recognize this problem, which is good. The problem may be dealt with in different ways. The authors used the hCG ratio but make no reference to this method. While the method has been used by others, it is not entirely clear how it adjusts for the differences in sampling days. According to the data provided, the ratios in the biopsy group and the reference group (ICSI) are significantly different, but this differences is not taken into account in the subsequent analyses. No discussion of this issue is provided.

Forth, the data set had missing data (as is invariably the case in such studies). The authors used mean imputation rather than multiple imputation. No arguments are given, but while multiple imputation is supposed to reduce the risk of bias (due to non-random selection), mean imputation may not do the trick. A thought on the choice og imputation method would have been appreciated.

Fifth, the selection problem becomes evident when comparing the unadjusted results with the results from the mathced analyses. According to the adjusted analyses in figure 3, the biopsy group was at significantly increased risk og GDM and umbilical cord abnormalities compared to the ICSI group. An asterisk marks the two analyses as ‘statistically significant’. But in table 3, the biopsy group is only significantly associated with GDM, while no p-value is provided for umbilical cord abnormalities (presumably because of no events in the ICSI group). Such descrepencies makes one wonder how the results in fugures and tables actually compare?

Finally, if GDM was the only outcome significantly associated with the exposure, may this simply be due to random error? Considering the number of statistical tests perfomed, the risk of type 1 error is considerable. A discussion of the risk of

type 1 error is, unfortunately, rarely seen.