Article details: 2018-0046	
Title	**Trends in prenatal testing service utilization for fetal chromosomal aneuploidies in Ontario: A descriptive study**
Authors	Tianhua Huang PhD, Shelley Dougan MPA MSc, Mark Walker MD, Christine M. Armour MD, Nan Okun MD
Reviewer 1	Lee Shulman MD
Institution	Ob/GYN, Northwestern University, Chigago, Ill.
General comments (author response in bold)	Comments to the Author This is a well written paper describing the outcomes of a retrospective analysis of a large data set of prenatal screening and diagnostic outcomes in Ontario. While the basis for this paper is not new or novel with many such papers being published over the course of the past 7 years of availability of NIPT, the uniqueness and importance of this paper is the large data set used for analysis as well as the clear recognition by the authors of the importance of public policy in determining prenatal screening and diagnostic choices. My only recommendation would be that the authors consider to further emphasize the seminal role of public policy (i.e., who pays for the test) in guiding the choices of women seeking information about their pregnancies. This paper presents important information of interest to our readers and to clinicians and health care administrators worldwide. **The impact of public funding of cfDNA screening was further discussed in Interpretation.**
Reviewer 2	Douglas Wilson MD
Institution	Obstetrics and Gynecology, University of Calgary, Calgary, Alta.
General comments (author response in bold)	Comments to the Author 1. Well done comprehensive review. 2, page 5 line 81 cf DNA. **Corrected.** 3. Results page 10 line 167 add July 1, 2012 - March 31, 2016. **Added.** 4. Results : are you able to comment on cf DNA failures / no call and the number of no result CVS / amnio 5. Results: are you able to comment on the birth prevalence of the screened for aneuploidies in the provincial areas either due to not screened or failed screening due to the variable uptake across the areas. **This paper is focused on utilization of prenatal screening and testing. We plan to report on no call result of cfDNA, factors associated with no call result and impact of cfDNA screening on the prevalence of aneuploidies in separate studies.** 6. Results: are you able to comment on the procedure loss rates for CVS / amnio 7. Overall: very well done. **BORN collects information on all prenatal diagnostic procedures and pregnancy outcomes of all live births and stillbirths. Whether a pregnancy with missing outcome in BORN was a miscarriage would need further investigation. We are not able to report procedure related loss rates in this paper.**
Reviewer 3	Anne-Marie Laberge MD PhD
Institution	Faculte de medicine, Université de Montréal, Montréal, Que.
General comments (author response in bold)	Comments to the Author This interesting study provides relevant population-level data to inform current and future policy on prenatal screening and diagnostic testing for fetal aneuploidies. The methods used have specific limitations, including potential ascertainment bias that should be better addressed in the discussion. Potential ascertainment bias was discussed in Interpretation (under limitations, line 333-346). This manuscript has multiple typos, which may be due to multiple revisions, but should have been caught before submission. **See detailed comments in file.** This interesting study provides relevant population-level data to inform current and future policy on prenatal screening and diagnostic testing for fetal aneuploidies. The methods used have specific limitations, including potential ascertainment bias that should be better addressed in the discussion.). Major comments: Line 74: please explain what « enhanced » FTS means. **Added an explanation about enhanced FTS** Line 75: DR and FPR are not the same for IPS and FTS. Please distinguish between the two. Okun et al (reference 2) describes DR as 88.4% for IPS and 83.9% for FTS. A 4.5% difference in DR at a population level (i.e. over tens of thousands of pregnancies) makes a difference. Also, Okun et al describe « positive rates » not FPR. **Rephrased the sentence. For simplicity, detection rates and false positive rates of IPS and FTS were not compared in the paper as both tests were discontinued in Ontario. References on the performance of different multiple marker screening tests were provided.** In Datasets, there seems to be a potential ascertainment bias. First, why is data from one cytogenetic laboratory not available? What proportion of pregnancies does that represent? Is that cytogenetic laboratory somehow different in terms of population tested than the others? Second, authors mention that unmatched records are most likely from miscarriages or terminations before 20 weeks of gestation. It is not explicitly stated whether these unmatched records were excluded of the analysis or not. If they were excluded, this could bias results: pregnancies that end in miscarriage or termination before 20 weeks are more likely to be due to fetal aneuploidy, and therefore may be more likely to have used MMS, cfDNA and/or PND, especially if fetal anomalies were present. This also means that all estimates

about uptake apply to uptake in pregnancies that continued to term (or at least to third trimester?). It is unclear from the information available in the manuscript what the inclusion criteria are for BORN. Are stillbirths included? How far along in the pregnancy is the pregnancy included? This affects how the final results about uptake can be interpreted and how they can be extrapolated to other populations.
In the analysis, the numerator and denominator are defined as « singleton pregnancies », but it should be specified which pregnancies are included vs excluded (e.g. only full term? Live born?). See comment about datasets above.
Added details about datasets (such as unmatched records, miscarriages and terminations and missing data) in 'Methods' (line 145-161) and discussed limitations including potential bias in 'Interpretation' (line 333-346).

Discussion should address the limitations raised in the two points above about inclusion and exclusion criteria of pregnancies, as it could lead to ascertainment bias.
Discussed limitations including potential bias in 'Interpretation'.

Discussion should address wide range of uptake by region, especially since two figures focus on uptake by region. This is briefly addressed (lines 260-264) but does not discuss the study's results, just what is known in the literature. The authors should raise questions or hypotheses based on their results. If the authors do not want to discuss these points in more detail, figures 1 and 2 become much less useful: they raise questions for the reader that are not explored by the authors.
Hypotheses and possible reasons of regional variations were discussed in Interpretation (line 293-309).

Line 219: « There was a small increase in the overall uptake of MMS… » When? Compared to what previous time period?
The uptake rates of MMS before and after funding were given in figure 3 and the second paragraph of 'Results'.

Line 236-244: The guidelines cited as « previous guidelines » that « continue to be operative » were the guidelines that were in effect at the time of the pregnancies (2012-2016), since new Canadian guidelines only became available in 2017. It is to be expected that they were still operative. The availability of a new test, especially if reimbursed only in high risk women, would not be expected to change how risk is assessed, i.e. through MMS. « Delayed incorporation » of new technologies may also have to do with access issues, especially in remote areas. This point needs to be nuanced or clarified.
The section has been revised for clarification.

Line 253: Be careful with the use of the terms « intermediate » or« lower » risk groups in this context. These are all women at high risk, even if some women are at even higher risk than others. I suggest you find a different way of expressing this idea.
Changed the wording to specify risk range rather than using terms 'intermediate' or 'low' risks.

Figure 3: Should use cfDNA, since that is the term used in the rest of the manuscript. Also, why is there no cfDNA pre-funding? The text mentions small percentages of women using cfDNA pre-funding. Finally, denominators for percentages need to be clearer. It may be better to put all numerators on same denominator.
Denominators for percentage were added. We have kept the current denominators as we felt it is more informative to compare PND among MMS positive women before and after funding rather than among all pregnancies. The number of cfNDA screening is very small pre-funding. We excluded it from the figure for simplicity as the main focus of the figure is the change in the utilization of PND.

Figure 4: All uptake values should be on the same scale. It is misleading right now, because cfDNA looks almost as frequent as MMS. The three bottom lines will be close together, but will be clearer to the reader.
We have removed the line about MMS and using a single Y axis for this figure.

Table 1:
- Title should say « by level of risk based on MMS», not « by the risk group of MMS »
- Third column is also percentages (not #).
- One decimal place would be enough in this context, but all values should be at least formated with the same number of decimal places.

Changes have been made accordingly.

Minor comments:
Lines 69-70: To what time period does this statement refer? Is that prior to 2012? In 2012? Please specify.
Added date range.

Lines 70-72: Following sentence needs to be revised: « a ultrasound marker nuchal translucency (NT) measured between 11 and 13 weeks of gestation ». Should read « an ultrasound marker, nuchal translucency (NT), … »
Corrected.

Line 81: « CfNDA » should be corrected to read « CfDNA »
Corrected.

Lines 94-95: Please quantify FPR and FNR to support claim that they are « low ». Even if they are lower than for IPS or FTS, they are still not low enough for CfDNA to be considered diagnostic. Statement should contextualize these characteristics and reflect that the point is that they are lower than with previously available screening.
CfDNA is not used as a diagnostic test. It can be used as a secondary screening test to reduce the number need for

	diagnostic testing. The performance of cfDNA has been given at in previous paragraph in the 'Introduction'. Line 180: delete « with » (« there were 10,312 … ») **Corrected.** Line 182: « uptake » instead of « update » **Corrected.** Line 186: Add a comma after 2014: « After women… in 2014, the uptake of… » **Added a comma.**