Journal of Computer Science and Cybernetics, V.40, N.2 (2024), 103–115
DOI no. 10.15625/1813-9663/20014

BENCHMARKING NIPT ALGORITHMS ON DETECTING NUMERICAL CHROMOSOME TRISOMY

LE SY VINH

University of Engineering and Technology, Vietnam National University Hanoi,
E3 Building, 144 Xuan Thuy, Cau Giay, Hanoi 10000, Viet Nam

Abstract. Noninvasive prenatal test (NIPT) is a widely used screening method to detect trisomy on chromosomes 13, 18, and 21. The lack of positive samples prevents us from examining the performance of NIPT algorithms on detecting trisomy on other chromosomes. Recently, we have introduced an efficient computational method to generate positive samples with trisomy from negative samples. In this paper, we applied the simulation method to generate 4600 positive samples for all 22 autosomal chromosomes as well as the X chromosome in females; and reused 1250 negative samples to assess the performance of algorithms CNVKit, WisecondorX, and VINIPT in detecting numerical chromosome aberrations. Experiments showed that WisecondorX had a sensitivity of 99.95% and a specificity of 97.2% on determining trisomy aberrations. VINIPT could detect all positive samples (i.e., sensitivity of 100%) and correctly determined 99.4% negative samples (i.e., specificity of 99.4%). The CNVkit algorithm was not as accurate as the WisecondorX and VINIPT algorithms. Its performance on some chromosomes such as chromosome 19 needs to be improved. WisecondorX and VINIPT could serve as reliable tools for analyzing NIPT data.

Keywords. Non-invasive prenatal test (NIPT); WisecondorX; VINIPT; CNVKit; Autosomal trisomy; Simulation.

1. INTRODUCTION

A normal human genome has 23 pairs of chromosomes including 22 pairs of autosomal chromosomes and one pair of sex chromosomes. Females have two X chromosomes while males have one X chromosome and one Y chromosome. Numerical chromosome aberrations (i.e., losing one copy or having one additional copy of a chromosome) might occur in fetal genomes during the pregnancy resulting in different serious disorders. The most prevalent numerical chromosome aberrations are trisomy on chromosomes (i.e., having an extra copy of a chromosome) including well-known notorious trisomy on chromosome 13 (T13 or Patau syndrome), chromosome 18 (T18 or Edwards syndrome), and chromosome 21 (T21 or Down syndrome).

Detecting numerical chromosome aberrations during the first weeks of pregnancy is one of the most crucial tests for pregnant women. Non-invasive prenatal test (NIPT) has been developed and widely used in practice to detect trisomy on chromosomes 13, 18, and 21.

Corresponding author.
E-mail addresses: vinhls@vnu.edu.vn (L.S. Vinh).

© 2024 Vietnam Academy of Science & Technology

To undertake the NIPT, cell-free DNA (cfDNA) is extracted from the blood of the mother as early as the 10^{th} week which includes a mixture of both maternal and fetal genomes. Statistical methods have been developed to detect numerical chromosome aberrations in the fetal genome from its cfDNA in the maternal blood such as NIFTY [3], Wisecondor [12], its improvement WisecondorX [11], and CNVkit [13].

The NIPT algorithms have a high level of sensitivity and specificity in detecting trisomy on chromosomes 13, 18, and 21 [7, 10, 15]. As almost all NIPT samples are normal (negative), a large number of negative samples are available from genetic testing centers. The negative samples can be used to measure the specificity (the ability to detect negative samples) of NIPT algorithms. However, the sensitivity (the ability to detect positive sample) of the NIPT algorithms is not easily evaluated due to the lack of positive samples publicly available to the scientific community.

Simulating positive samples is a good approach to assess the sensitivity of NIPT methods. Recently, we have proposed a simple computational approach to simulate positive samples with aberrations on autosomal chromosomes from negative samples [8]. We simulated positive samples with trisomy on chromosomes 13, 18, and 21 to preliminarily evaluate the sensitivity of NIPT software WisecondorX and VINIPT. In this paper, we applied the simulation method to create positive samples with trisomy on all 22 autosomal chromosomes as well as X chromosome in females. We used the positive samples and 1250 negative samples to measure the sensitivity and specificity of three algorithms CNVkit, WisecondorX, and VINIPT.

2. DATA AND METHODS

2.1. Data

We reused negative samples from singleton pregnancies in our recent study [8]. The samples were sequenced by the MGISEQ-200 platform with 50 bp single-end reads. We selected 200 negative female samples with the number of reads from 5 to 15 million, and the fetal DNA fraction of at least 4.0 from the dataset to create positive samples. We created 200 positive samples with trisomy for each chromosome (i.e., chromosome 1 to chromosome 22 and chromosome X). In this study, we simulated 200 positive samples for each chromosome to avoid computational overload. In total, we created 4600 positive samples (200 samples $\times$ 23 chromosomes) for measuring the sensitivity of NIPT algorithms. Their fetal DNA fractions range from 4.1% to 20.9% with a mean of about 10%. The average number of reads per sample is about 9.7 million ($\sim$ 0.15x). We also selected 1250 negative samples with at least 2.5 million reads to examine the specificity of the algorithms. The mean fetal DNA fraction of the negative samples is about 9.7% (ranging from 3.4% to 21%). The negative samples contain from 2.7 to 35.9 million reads with a mean of 11.6 million reads ($\sim$ 0.17x). The statistics of the samples are summarized in Table 1.

2.2. NIPT procedure

Non-invasive prenatal test (NIPT) provides a safe and efficient method for screening aberrations in the genomes of developing fetuses. To carry out NIPT, blood from the mother is collected as early as the 10th week of pregnancy, and samples are subject to a series of both wet-lab and bioinformatics steps (see Figure 1) to determine numerical chromosome

Table 1: The statistics of fetal DNA fraction and the number of reads of negative and simulated positive NIPT samples.

		Fetal DNA fraction (%)				#Reads (million)			
Dataset	#Samples	Min	Mean	Std	Max	Min	Mean	Std	Max
Simulated positive samples	4600	4.1	10.0	2.5	20.9	5.4	9.7	1.5	14.8
Negative samples	1250	3.4	9.7	2.9	21.0	2.7	11.6	6.3	35.5

aberrations. The bioinformatics steps including NIPT algorithms play a pivotal role in the analysis and interpretation of genomic aberrations originating from the fetal genome.

Wet-lab steps: They include blood collection from pregnant women, cfDNA extraction from the maternal blood, and genome sequencing.

- Blood collection: The blood from pregnant women contains a mixture of cfDNA from genomes of both mother and developing fetus. A small amount of blood is collected from the pregnant woman, typically as early as the 10^{th} week of pregnancy.
- cfDNA extraction: The cfDNA from the maternal blood sample is extracted. Most of cfDNA ($\sim 90\%$) originate from the maternal genome and a small portion ($\sim 10\%$) are from the fetal genome.
- Genome sequencing: The cfDNA is sequenced by high-throughput sequencing platforms such as Illumina. The step results in multi-million short reads. In practice, the read coverage per sample typically ranges from 0.1x ($\sim$ 6 million reads of 50 bps in length) to 1x.

The Bioinformatics phase consists of three key steps: data processing, fetal DNA fraction estimation, and identification of numerical chromosome aberrations using NIPT algorithms.

Data processing: The sequenced short reads undergo several data processing steps including sequence alignment and sequence cleaning. Mapping short reads to the reference genome to determine their positions in the genome using a genome aligner such as Bowtie2 [5] or BWA [6]. To obtain high-quality data for downstream analyses, the short reads ambiguously mapped to multiple positions might be removed; and duplications should be also eliminated except one.

Fetal DNA fraction estimation: Applying computational methods to estimate the fetal DNA fraction. If the fetal DNA fraction is not enough (e.g., smaller than 5%), we have to re-collect blood and re-sequence data to obtain a higher fetal DNA fraction (the longer pregnancy time leads to higher fetal DNA fraction). Technically, the fetal DNA fraction can be estimated by computational methods [14] such as SeqFF [4] that can efficiently work for both male and female samples.

NIPT algorithms: A number of NIPT algorithms have been proposed to detect numerical chromosome aberrations such as NIFTY [3], CNVkit [13], Wisecondor [12], and its improvement WisecondorX [11]. If the fetus of a test sample has an additional copy of chromosome h, the read percentage of chromosome h in the test sample will increase in comparison with that in normal samples. Detecting the change in a fetal genome is a challenging problem because fetal DNA only makes up a small part of the NIPT data. The NIPT algorithms typically employ statistical methods and the read coverage information to detect numerical chromosome aberrations in the fetal genome. The general idea of NIPT

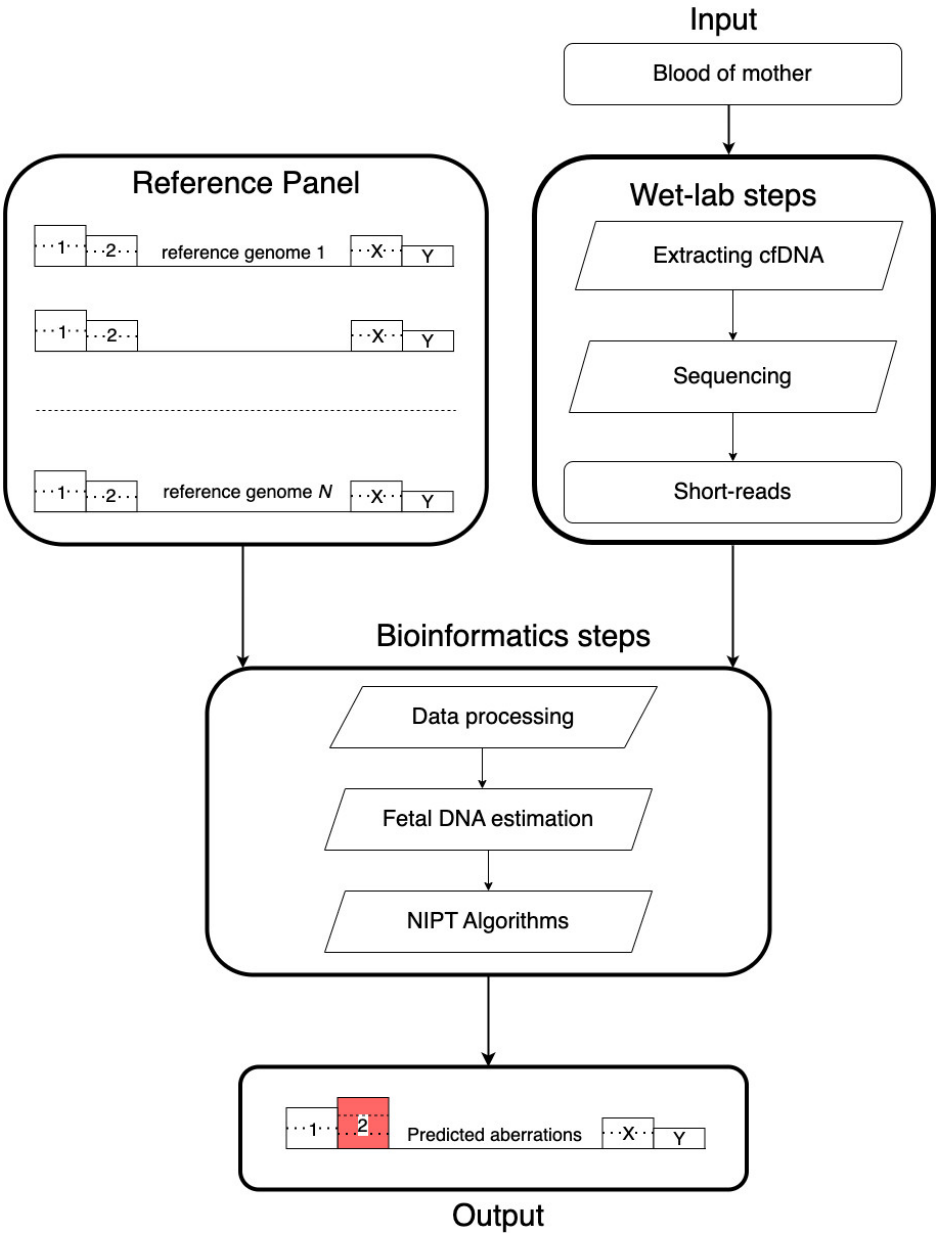

Figure 1: The NIPT procedure to detect numerical chromosome aberrations from the maternal blood. The procedure includes both wet lab and bioinformatics steps.

algorithms is to compare the read coverage in the test sample with those in the normal samples to determine numerical chromosome aberrations in the test sample.

Numerical chromosome aberrations are determined from aberrations on small regions (called bins) in chromosomes. A genome is normally divided into equal-size bins with lengths of thousands to millions of base pairs such that each bin contains a sufficient number of reads. The read coverage percentages of bins in normal samples are analyzed to construct a reference

panel and used as the basis to compare with that in the test sample. If the read coverage of bin b in the test sample considerably deviates from that in the reference panel, bin b is considered as abnormal. The NIPT algorithms use different statistical tests to assign a z-score for each bin. The overall z-score of a chromosome is estimated from the z-scores of its bins. A chromosome is predicted abnormal if its overall z-score exceeds a predefined z-score threshold.

Comparing the read coverages of bins in a test sample with that in normal samples requires that the test sample and reference samples are sequenced under the same conditions. This requirement might not be completely fulfilled in practice. To overcome the strictly technical requirement, within-sample copy number aberration detection algorithms Wisecondor [12] and its improvement WisecondorX [11] have been proposed. For each bin b in chromosome h, the within-sample algorithm determines a set of reference bins in other chromosomes that have similar read coverages to b in normal samples. This strategy guarantees that bin b and its reference bins in the test sample are sequenced under the same condition. The read coverage of b in the test sample will be compared with that of its reference bins to assign a z-score for bin b. Stouffer's z-score sliding window approach or circular binary segmentation algorithm [9] can be used to determine aberrations on chromosomes from z-scores of bins.

2.3. Positive sample simulation

The trisomy on chromosomes occurs when the fetal genome has an additional copy of a chromosome. Note that we are unable to separate the cfDNA of the maternal genome from that of the fetal genome. Although computational methods have been proposed to simulate short reads, they are not designed for NIPT data with such complex characteristics as cfDNA artifacts, differences in fragment size, and quality of short-reads from maternal and fetal genomes.

Recently, we proposed a computational method to create positive samples from negative samples without simulating new short reads [8]. Let Q be a normal NIPT sample in which each autosomal chromosome has two copies. Sample Q has trisomy on chromosome h if the fetal genome contains three copies of chromosome h. Given the NIPT data of sample Q, adding new short reads to create an additional copy of a chromosome h in the fetal genome will increase the read percentage of chromosome h in comparison with those of other chromosomes. Equivalently, we do not change the read percentage of chromosome h, but decrease the read percentage of other chromosomes by deleting their short reads.

Let f be the fraction of fetal DNA in maternal blood, i.e., $f\%$ of short reads come from the fetal genome, and $(100 - f)\%$ of short reads are from the maternal genome. The read percentage C_h of chromosome h will be increased to $C_h = C_h(1 + 0.5 \times f)$ when inserting an additional copy of chromosome h in the fetal genome. This is equivalent to keep C_h unchanged and decrease the read coverage C_i of other chromosomes $i (i \neq h)$ to $C_i' = \frac{C_i}{(1+0.5\times f)}$.

3. RESULTS

We assessed the sensitivity and specificity of three algorithms CNVKit, WisecondorX, and its improvement VINIPT [8]. First, we reported the performance of the algorithms with

the z-score cutoff threshold of 3.3 as used in [8]. We measured the sensitivity of the NIPT algorithms on 4600 positive samples and their specificity on 1250 negative samples. We also investigated the performance of the algorithms with higher z-score cutoff thresholds of 5.0 and 7.0. We note that some other well-known algorithms such as NIFTY are commercialized and thus not publicly available for testing.

3.1. Sensitivity evaluation

First, we examined the sensitivity of CNVKit, WisecondorX, and VINIPT algorithms on 4600 positive samples, i.e., 200 positive samples for each chromosome. Table 2 shows the prediction results on all chromosomes using the z-score cutoff threshold of 3.3 for the three NIPT algorithms.

Table 2: The sensitivity of WisecondorX, VINIPT, and CNVKit algorithms with the z-score cut-off threshold of 3.3. The results are summarized from 4600 positive samples with trisomy aberrations.

Chromosome	WisecondorX		VINIPT		CNVKit	
	#False negative	Sensitivity	#False negative	Sensitivity	#False negative	Sensitivity
Chr1	0	100%	0	100%	**2**	**99%**
Chr2	0	100%	0	100%	0	100%
Chr3	0	100%	0	100%	0	100%
Chr4	0	100%	0	100%	0	100%
Chr5	0	100%	0	100%	0	100%
Chr6	0	100%	0	100%	0	100%
Chr7	0	100%	0	100%	0	100%
Chr8	0	100%	0	100%	0	100%
Chr9	0	100%	0	100%	**28**	**86.2%**
Chr10	0	100%	0	100%	0	100%
Chr11	0	100%	0	100%	0	100%
Chr12	0	100%	0	100%	0	100%
Chr13	0	100%	0	100%	0	100%
Chr14	0	100%	0	100%	0	100%
Chr15	0	100%	0	100%	**8**	**96%**
Chr16	0	100%	0	100%	**6**	**97%**
Chr17	0	100%	0	100%	0	100%
Chr18	0	100%	0	100%	0	100%
Chr19	0	100%	0	100%	**9**	**95.5%**
Chr20	0	100%	0	100%	0	100%
Chr21	0	100%	0	100%	48	76%
Chr22	**1**	**99.5%**	0	100%	10	95%
ChrX	0	100%	0	100%	**1**	**99.5%**
All 23 chromosomes	**1**	**99.95%**	**0**	**100%**	**112**	**97.56%**

The WisecondorX was able to detect almost all positive samples except one sample with aberration on chromosome 22, i.e., its sensitivity is equivalent to 99.95%. The VINIPT could detect all positive samples, i.e., the sensitivity of 100%. The CNVKit falsely predicted 112 positive samples as negative, e.g., 28 cases on chromosome 9 or 48 cases on chromosome 21. The overall sensitivity of CNVKit in detecting positive trisomy samples was 97.56%.

Figure 2 shows the z-score distributions of WisecondorX, VINIPT, and CNVKit algorithms on positive samples. Almost all positive samples have high z-scores, except 112 false negative samples predicted by CNVKit that have low z-scores (displayed as outliers). The average z-score predicted by the CNVkit algorithm is lower than that by the WisecondorX and VINIPT algorithms.

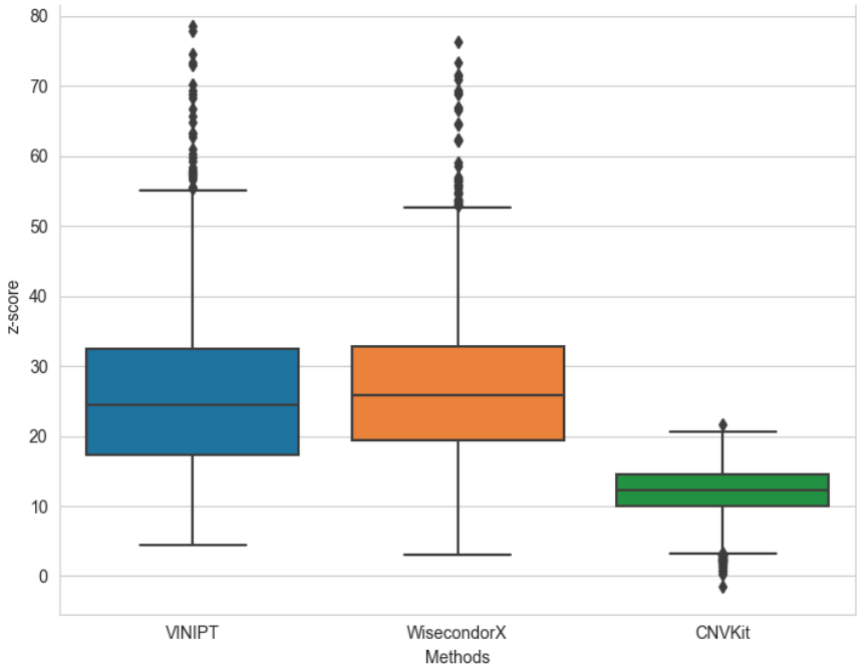

Figure 2: The distributions of z-scores from VINIPT, WisecondorX, and CNVKit methods on positive samples with trisomy aberrations.

3.2. Specificity evaluation

We evaluated the specificity of WisecondorX, VINIPT, and CNVKit algorithms on 1250 negative samples. Table 3 summarizes the specificity of the algorithms for 23 chromosomes. The WisecondorX falsely assigned 187 out of 1250 negative samples to positive samples, i.e., its specificity equals 85.1%. The VINIPT could determine almost all negative samples (it wrongly predicted 8 false positive samples) and has a specificity of 99.4%. The CNVKit incorrectly predicted 323 negative samples as positive samples, i.e., its specificity of only 74.2%. We observed that a large number of false positive samples was on chromosome 19 (26 from WisecondorX; 3 from VINIPT; and 194 from CNVKit). The chromosome 19 is a difficult chromosome for all NIPT algorithms due to its rich GC content and high content of hypermutable CpG sites that make chromosome 19 highly diverse [1].

The distributions of z-scores derived from VINIPT, WisecondorX, and CNVkit methods on negative samples are illustrated in Figure 3. We see that outlier samples (false positive samples) have z-scores greater than the predefined z-score threshold of 3.3.

3.3. Z-score threshold evaluation

Changing the z-score cut-off threshold will change the sensitivity and specificity of NIPT algorithms. Increasing the z-score threshold helps reduce the false positive rate (i.e., increasing the specificity), but raises the false negative rate (i.e., reducing the sensitivity). In contrast, decreasing the z-score threshold might increase the sensitivity, but reduce the speci-

Table 3: The specificity of WisecondorX, VINIPT, and CNVKit algorithms with the z-score cut-off threshold of 3.3. The results are summarized from 1250 negative samples.

Chromosome	WisecondorX		VINIPT		CNVKit	
	#False positive	Specificity	#False positive	Specificity	#False positive	Specificity
Chr1	4	99,7%	0	100%	2	99,8%
Chr2	4	99,7%	0	100%	1	99,9%
Chr3	5	99,6%	0	100%	9	99,3%
Chr4	5	99,6%	0	100%	20	98,4%
Chr5	8	99,4%	0	100%	11	99,1%
Chr6	11	99,1%	0	100%	8	99,4%
Chr7	10	99,2%	1	99,9%	7	99,4%
Chr8	13	99%	0	100%	6	99,5%
Chr9	2	99,8%	0	100%	26	97,9%
Chr10	8	99,4%	1	99,9%	1	99,9%
Chr11	4	99,7%	0	100%	0	100%
Chr12	4	99,7%	0	100%	1	99,9%
Chr13	4	99,7%	0	100%	22	98,2%
Chr14	4	99,7%	0	100%	1	99,9%
Chr15	13	99%	0	100%	1	99,9%
Chr16	6	99,5%	1	99,9%	19	98,5%
Chr17	7	99,4%	0	100%	0	100%
Chr18	7	99,4%	0	100%	10	99,2%
Chr19	26	97,9%	3	99,8%	194	84,5%
Chr20	6	99,5%	0	100%	2	100%
Chr21	6	99,5%	0	100%	0	100%
Chr22	4	99,7%	2	99,8%	1	100%
ChrX	18	98,6%	0	100%	27	97,8%
All 23 chromosomes	**187**	**85.1%**	**8**	**99,4%**	**323**	**74.2%**

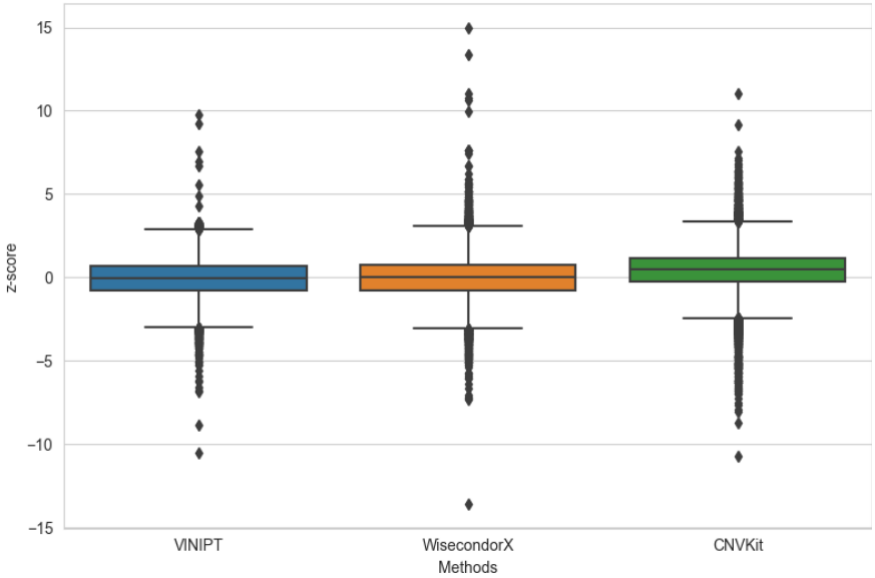

Figure 3: The distributions of z-scores from VINIPT, WisecondorX, and CNVkit methods on negative samples.

ficity. We examined the performance of WisecondorX, VINIPT, and CNVkit with higher z-score cutoff thresholds of 5.0, and 7.0. Table 4 shows the overall sensitivity and specificity of the NIPT algorithms with the z-score thresholds of 3.3, 5.0 and 7.0. The sensitivity and specificity of the NIPT algorithms for every chromosome for z-scores of 5.0 and 7.0 are summarized in Table S1-S4 in the Appendix.

WisecondorX has a sensitivity of 99.95% for both z-score thresholds of 3.3 and 5.0; and a lower sensitivity of 99.9% with a z-score threshold of 7.0. Its specificity gets much higher from 85.1% to 99.2% when increasing the z-score threshold from 3.3 to 7.0. The sensitivity and specificity of VINIPT slightly changed when alternating the z-score thresholds from 3.3 to 7.0, i.e., the sensitivity decreased from 100% to 99.5% while the specificity increased from 99.4% to 99.8%.

The specificity of CNVKit was significantly better (from 74.2% to 94.9%) when using the z-score threshold of 5.0 instead of 3.3, while its sensitivity reduced from 97.56% to a lower level of 93.54%. The CNVKit with a z-score threshold of 7.0 had an overall sensitivity of 86.36%; and considerably low sensitivity on some chromosomes such as only 7.5% on chromosome 21. It is challenging for CNVKit to have a high specificity on chromosome 9 possibly due to its highly structurally polymorphic and many intra and inter-chromosomal duplications [2].

Table 4: The sensitivity and specificity of WisecondorX, VINIPT, and CNVkit algorithms with different z-score cut-off thresholds.

z-score thresholds	WisecondorX		VINIPT		CNVkit	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
3.3	99.95%	85.1%	100%	99.4%	97.56%	74.2%
5.0	99.95%	97.7%	99.91%	99.5%	93.54%	94.9%
7.0	99.9%	99.2%	99.5%	99.8%	86.36%	99.6%

4. DISCUSSION

The NIPT method has been widely used to screen numerical aberrations on chromosomes due to its safe (non-invasive), cheapness, and ease of implement on practice. The performance of NIPT algorithms has been evaluated for some chromosomes related to popular disorders, i.e., Down syndrome, Patau syndrome, and Edwards syndrome. The specificity of NIPT algorithms can be measured from a large number of negative samples, however, the sensitivity is not easily established due to the lack of positive samples.

We applied a simulation method to create 4600 positive samples; and re-used 1250 negative samples to assess the performance of well-known NIPT algorithms CNVKit and WisecondorX (and its improvement VINIPT). Some other NIPT algorithms are implemented in commercial software such as NIFTY and are not publicly available for testing. Experiments showed that WisecondorX and VINIPT had a high level of sensitivity in detecting trisomy on all chromosomes, and VINIPT performed better than WisecondorX in terms of specificity. The CNVKit had lower sensitivity and specificity than those of WisecondorX and VINIPT. Its specificity needs to be improved, especially for difficult chromosomes such as chromosome 19.

The z-score cutoff threshold affects the sensitivity and specificity of NIPT algorithms. Experiments showed that VINIPT worked well with the z-score cutoff of 3.3 by reaching a

sensitivity of 100% and specificity of 99.4%. VINIPT with the z-score threshold of 5.0 or 7.0 resulted in a number of false negative samples. On the other hand, the z-score threshold of 5.0 was a good choice for both WisecondorX and CNVKit because it helped significantly reduce the false positive rates of the two algorithms. We note that using a high z-score threshold will reduce the sensitivity of all NIPT algorithms.

Although NIPT has been developed for decades, numerical chromosome aberrations detected by the NIPT algorithms might be false positive. Therefore, aberrations predicted by NIPT must be verified by other specialized prenatal tests such as karyotype. A positive prediction from NIPT should be carefully explained to pregnant women to avoid any unnecessary termination.

ACKNOWLEDGMENT

The technical supports from Nguyen Huu Tinh and Nguyen Van Huynh are highly appreciated.

REFERENCES

[1] R. A. Harris, M. Raveendran, K. C. Worley, and J. Rogers, "Unusual sequence characteristics of human chromosome 19 are conserved across 11 nonhuman primates," *BMC Evolutionary Biology*, vol. 20, no. 1, Feb. 2020. [Online]. Available: http://dx.doi.org/10.1186/s12862-020-1595-9

[2] S. J. Humphray *et al.*, "DNA sequence and analysis of human chromosome 9," *Nature*, vol. 429, no. 6990, p. 369–374, May 2004. [Online]. Available: http://dx.doi.org/10.1038/nature02465

[3] F. Jiang *et al.*, "Noninvasive fetal trisomy (NIFTY) test: An advanced noninvasive prenatal diagnosis methodology for fetal autosomal and sex chromosomal aneuploidies," *BMC Medical Genomics*, vol. 5, no. 1, Dec. 2012.

[4] S. K. Kim *et al.*, "Determination of fetal DNA fraction from the plasma of pregnant women using sequence read counts: Determination of fetal DNA fraction from the plasma of pregnant women using sequence read counts," *Prenatal Diagnosis*, vol. 35, no. 8, p. 810–815, Jun. 2015. [Online]. Available: http://dx.doi.org/10.1002/pd.4615

[5] B. Langmead and S. L. Salzberg, "Fast gapped-read alignment with bowtie 2," *Nature Methods*, vol. 9, no. 4, p. 357–359, Mar. 2012. [Online]. Available: http://dx.doi.org/10.1038/nmeth.1923

[6] H. Li and R. Durbin, "Fast and accurate long-read alignment with burrows–wheeler transform," *Bioinformatics*, vol. 26, no. 5, p. 589–595, Jan. 2010. [Online]. Available: http://dx.doi.org/10.1093/bioinformatics/btp698

[7] T. Mokveld, Z. Al-Ars, E. A. Sistermans, and M. Reinders, "A comprehensive performance analysis of sequence-based within-sample testing nipt methods," *PLOS ONE*, vol. 18, no. 4, p. e0284493, Apr. 2023. [Online]. Available: http://dx.doi.org/10.1371/journal.pone.0284493

[8] T. Nguyen, H. Nguyen, M. Pham, and V. Le, "An efficient computational method to create positive nipt samples with autosomal trisomy," in *The 12th International Conference on Bioinformatics and Computational Biology*, Tokyo, 2024.

[9] A. B. Olshen, E. S. Venkatraman, R. Lucito, and M. Wigler, "Circular binary segmentation for the analysis of array-based DNA copy number data," *Biostatistics*, vol. 5, no. 4, p. 557–572, Oct. 2004. [Online]. Available: http://dx.doi.org/10.1093/biostatistics/kxh008

[10] Paluoja *et al.*, "Systematic evaluation of NIPT aneuploidy detection software tools with clinically validated NIPT samples," *PLOS Computational Biology*, vol. 17, no. 12, p. e1009684, Dec. 2021. [Online]. Available: http://dx.doi.org/10.1371/journal.pcbi.1009684

[11] L. Raman, A. Dheedene, M. De Smet, J. Van Dorpe, and B. Menten, "WisecondorX: improved copy number detection for routine shallow whole-genome sequencing," *Nucleic Acids Research*, vol. 47, no. 4, p. 1605–1614, Dec. 2018. [Online]. Available: http://dx.doi.org/10.1093/nar/gky1263

[12] R. Straver, E. A. Sistermans, H. Holstege, A. Visser, C. B. M. Oudejans, and M. J. T. Reinders, "WISECONDOR: detection of fetal aberrations from shallow sequencing maternal plasma based on a within-sample comparison scheme," *Nucleic Acids Research*, vol. 42, no. 5, p. e31–e31, Oct. 2013. [Online]. Available: http://dx.doi.org/10.1093/nar/gkt992

[13] E. Talevich, A. H. Shain, T. Botton, and B. C. Bastian, "CNVkit: Genome-wide copy number detection and visualization from targeted dna sequencing," *PLOS Computational Biology*, vol. 12, no. 4, p. e1004873, Apr. 2016. [Online]. Available: http://dx.doi.org/10.1371/journal.pcbi.1004873

[14] van Beek *et al.*, "Comparing methods for fetal fraction determination and quality control of NIPT samples: Comparing fetal fraction and quality control tools for NIPT," *Prenatal Diagnosis*, vol. 37, no. 8, p. 769–773, Jul. 2017. [Online]. Available: http://dx.doi.org/10.1002/pd.5079

[15] H. Zhang *et al.*, "Non-invasive prenatal testing for trisomies 21, 18 and 13: Clinical experience from 146 958 pregnancies," *Ultrasound in Obstetrics amp; Gynecology*, vol. 45, no. 5, p. 530–538, Apr. 2015. [Online]. Available: http://dx.doi.org/10.1002/uog.14792

Received on January 26, 2024
Accepted on March 28, 2024

APPENDIX

Table S1: The sensitivity of WisecondorX, VINIPT, and CNVKit algorithms with the z-score cut-off threshold of 5.0. The results are summarized from 4600 positive samples with trisomy aberrations.

Chromosome	WisecondorX		VINIPT		CNVKit	
	#False negative	Sensitivity	#False negative	Sensitivity	#False negative	Sensitivity
Chr1	0	100%	0	100%	6	97%
Chr2	0	100%	0	100%	1	99,5%
Chr3	0	100%	0	100%	0	100%
Chr4	0	100%	0	100%	0	100%
Chr5	0	100%	0	100%	0	100%
Chr6	0	100%	0	100%	0	100%
Chr7	0	100%	0	100%	0	100%
Chr8	0	100%	0	100%	0	100%
Chr9	0	100%	0	100%	74	63%
Chr10	0	100%	0	100%	0	100%
Chr11	0	100%	0	100%	0	100%
Chr12	0	100%	0	100%	2	99%
Chr13	0	100%	0	100%	0	100%
Chr14	0	100%	0	100%	0	100%
Chr15	0	100%	0	100%	23	88,5%
Chr16	0	100%	0	100%	16	92%
Chr17	0	100%	0	100%	0	100%
Chr18	0	100%	0	100%	0	100%
Chr19	0	100%	2	99%	11	94,5%
Chr20	0	100%	0	100%	1	99.5%
Chr21	0	100%	1	99.5%	115	42.5%
Chr22	1	99.5%	1	99.5%	44	78%
ChrX	0	100%	0	100%	4	98%
All 23 chromosomes	**1**	**99.95%**	**4**	**99.91%**	**297**	**93.54%**

Table S2: The specificity of WisecondorX, VINIPT, and CNVKit algorithms with the z-score cut-off threshold of 5.0. The results are summarized from 1250 negative samples.

Chromosome	WisecondorX		VINIPT		CNVKit	
	#False positive	Specificity	#False positive	Specificity	#False positive	Specificity
Chr1	0	100%	0	100%	0	100%
Chr2	1	99,9%	0	100%	0	100%
Chr3	1	99,9%	0	100%	0	100%
Chr4	0	100%	0	100%	1	99,9%
Chr5	2	99,8%	0	100%	2	99,8%
Chr6	1	99,9%	0	100%	0	100%
Chr7	1	99,9%	1	99,9%	1	99,9%
Chr8	1	99,9%	0	100%	1	99,9%
Chr9	1	99,9%	0	100%	0	100%
Chr10	1	99,9%	0	100%	0	100%
Chr11	0	100%	0	100%	0	100%
Chr12	0	100%	0	100%	0	100%
Chr13	1	99,9%	0	100%	1	99,9%
Chr14	0	100%	0	100%	0	100%
Chr15	1	99,9%	0	100%	0	100%
Chr16	5	99,6%	1	99,9%	0	100%
Chr17	0	100%	0	100%	0	100%
Chr18	0	100%	0	100%	1	99,9%
Chr19	8	99,4%	3	99,8%	54	95,7%
Chr20	0	100%	0	100%	0	100%
Chr21	1	99,9%	0	100%	0	100%
Chr22	1	99,9%	1	99,9%	0	100%
ChrX	3	99,8%	0	100%	3	99,8%
All 23 chromosomes	**29**	**97,7%**	**6**	**99,5%**	**64**	**94,9%**

Table S3: The sensitivity of WisecondorX, VINIPT, and CNVKit algorithms with the z-score cut-off threshold of 7.0. The results are summarized from 4600 positive samples with trisomy aberrations.

Chromosome	WisecondorX		VINIPT		CNVKit	
	#False negative	Sensitivity	#False negative	Sensitivity	#False negative	Sensitivity
Chr1	0	100%	0	100%	17	91,5%
Chr2	0	100%	0	100%	1	99,5%
Chr3	0	100%	0	100%	0	100%
Chr4	0	100%	0	100%	0	100%
Chr5	0	100%	0	100%	0	100%
Chr6	0	100%	0	100%	0	100%
Chr7	0	100%	0	100%	0	100%
Chr8	0	100%	0	100%	0	100%
Chr9	0	100%	0	100%	129	35,5%
Chr10	0	100%	0	100%	2	99%
Chr11	0	100%	0	100%	2	99%
Chr12	0	100%	0	100%	14	93%
Chr13	0	100%	0	100%	1	99,5%
Chr14	0	100%	0	100%	2	99%
Chr15	0	100%	0	100%	55	72,5%
Chr16	0	100%	0	100%	46	77%
Chr17	0	100%	0	100%	4	98%
Chr18	0	100%	0	100%	1	99,5%
Chr19	2	100%	20	90%	23	88,5%
Chr20	0	100%	0	100%	16	92%
Chr21	1	100%	1	99,5%	185	7.5%
Chr22	3	100%	4	98%	141	29.5%
ChrX	0	100%	0	100%	4	98%
All 23 chromosomes	**6**	**99,9%**	**25**	**99,5%**	**627**	**86.36%**

Table S4: The specificity of WisecondorX, VINIPT, and CNVKit algorithms with the z-score cut-off threshold of 7.0. The results are summarized from 1250 negative samples.

Chromosome	WisecondorX		VINIPT		CNVKit	
	#False positive	Specificity	#False positive	Specificity	#False positive	Specificity
Chr1	0	100%	0	100%	0	100%
Chr2	0	100%	0	100%	0	100%
Chr3	0	100%	0	100%	0	100%
Chr4	0	100%	0	100%	0	100%
Chr5	2	99,8%	0	100%	0	100%
Chr6	0	100%	0	100%	0	100%
Chr7	0	100%	0	100%	0	100%
Chr8	1	99,9%	0	100%	1	99,9%
Chr9	0	100%	0	100%	0	100%
Chr10	0	100%	0	100%	0	100%
Chr11	0	100%	0	100%	0	100%
Chr12	0	100%	0	100%	0	100%
Chr13	0	100%	0	100%	0	100%
Chr14	0	100%	0	100%	0	100%
Chr15	0	100%	0	100%	0	100%
Chr16	1	99,9%	1	99,9%	0	100%
Chr17	0	100%	0	100%	0	100%
Chr18	0	100%	0	100%	0	100%
Chr19	4	99,7%	2	99,8%	2	99,8%
Chr20	0	100%	0	100%	0	100%
Chr21	0	100%	0	100%	0	100%
Chr22	1	99,9%	0	100%	0	100%
ChrX	1	99,9%	0	100%	2	99,8%
All 23 chromosomes	**10**	**99,2%**	**3**	**99,8%**	**5**	**99,6%**