Thesis, FRANZEN Rachelle

Auteur : Njonfang Tchuidjang, Ornela
Promoteur(s) : BOEMER, François; LUIS, Géraldine; Brysse, Anne
Faculté : Faculté de Médecine
Diplôme : Master en sciences biomédicales, à finalité spécialisée en assurance qualité
Année académique : 2025-2026
URI/URL : http://hdl.handle.net/2268.2/25972

Revalidation/Validation of the BabyDedect Newborn Screening Test By Next Generation Sequencing (NGS) in an ISO15189/IVDR context

Thesis prepared and submitted by: Ornéla NJONFANG TCHUIDJANG,

In partial fulfillment of the requirements for the master's degree in Biomedical Sciences, with a specialization in Quality assurance

Supervisors

Dr Phar. Biol.: François BOEMER

Dr. Sci : Luis Geraldine

Dr. Sci : Anne Brysse

Jury members

Dr. Sci Arnaud LAVERGNE

Dr. Sci Corinne FASQUELLE

Dr. Sci Laura MASSOZ

ACADEMIC YEAR: 2025 – 2026

ACKNOWLEDGMENTS

The work presented below was made possible thanks to the support, contribution, and participation of individuals to whom I would like to express my gratitude on this occasion.

First and foremost, I thank the Lord Jesus Christ, who is my ultimate material, financial, and intellectual provider.

Next, I thank my parents, Emmanuel NJONFANG and Hélène Dayong, as well as my brothers and sisters, Vanessa NJONFANG, Epiphanie NJONFANG, and Yvan NJONFANG, who have always provided intellectual, moral, and financial support, and who ensured the continuation of my studies from start to finish.

I would like to thank my supervisors: Dr.Phar.Biol François BOEMER, head of the Genetic Biochemistry Laboratory; Dr.Sci Geraldine LUIS, scientist at the Genetic Biochemistry Laboratory; and Dr.Sci Anne BRYSSE, scientist and quality manager of the Human Genetics Department. They were present during every step of my internship and the writing of this work, and found the time to answer my questions, concerns, and misunderstandings regarding quality assurance, the general operation of the laboratory, and scientific methodology.

I would also like to thank Dr.Sci Valerie JACQUEMIN, Dr.Sci Kristine HOVHANNESYAN, and technologist Myriam MNI, who were of invaluable help with the technical aspects of this work.

As my work also involves a considerable bioinformatics component, I would like to thank the entire bioinformatics team of the genetics department. In particular, I thank Dr. Leonor PALMEIRA, bioinformatics project manager in the Human Genetics Department, who always made herself available for discussions and explanations on bioinformatics and even biostatistical questions.

I cannot conclude this acknowledgments section without sending a special thank you to my life partner, Louis Christian MBEKWE, who, beyond intellectual or financial matters, is a true moral and spiritual support in all areas of my life.

Thank you to my cousins, especially the YOUSSEU family, Ange, Cedric, and Marcelle.

Thank you to everyone who, near or far, contributed to the realization of this work.

LIST OF ABBREVIATIONS

A,T,G,C,U	Adenine, Thymine, Guanine, Cytosine, Uracil
ACMG	American College of Medical Genetics and Genomics
AMP	Association for Molecular Pathology
BAM/CRAM	Binary alignment/Map - Compressed Reference-oriented Alignment Map
BELAC	Belgian accreditation
FMEA	Failure Mode and Effects Analysis
Bp,kb,Mb	Base pair, kilo base, Mega base
CNV	Copy Number Variation
DBS	Dry Blood Spot
DNA	Deoxyribonucleic acid
dNTP/ddNTP	deoxyribonucleotide triphosphates/dideoxynucleoside triphosphates
DRAGEN	Dynamic Read Analysis for GENomics
EQA	External quality Assessment
ELISA	Enzyme-Linked Immunosorbent Assay
FAMHP	Federal Agency for Medicines and Health Products
FN	False negative
FP	False positive
FPGA	Field Programmable Gate Array
GIAB	Genome In A Bottle
GIGA	Groupe Interdisciplinaire de Genoprotéomique Appliquée
GLIMS	Laboratory Information management system
HCR	High Confidence Region
Hg19	Human genome 19
INDELS	Insertions and Deletions
ISO	International Organization for Standardization
IVDR	In Vitro Diagnostic Regulation
LCMS	Liquid chromatography–mass spectrometry
LDT	Laboratory Developed Test
MDCG	Medical Device Coordination Group
MDR	Medical Device regulation
NBS	Newborn screening
NIHDI	National Institute for Health and Disability Insurance
ONE	Office de la Naissance et de l'Enfance
QA	Quality Assurance
IQC, IQR	Internal Quality Control; Interquartile range
QMS	Quality Management System
qPCR	Quantitative Polymerase Chain RE
RNA	Ribonucleic acid
SMA	Spinal Muscular Atrophy
SNP	Single Nucleotide Polymorphism
VAF	Variant Allele Frequency
VCF	Variant Call Format

WES	Whole Exome Sequencing
WGS	Whole genome Sequencing

LIST OF TABLES AND FIGURES

- **LIST OF TABLES**

- **LIST OF FIGURES**

ABSTRACT

Context: This work describes the BabyDetect neonatal screening genomic test for pediatric, treatable, genetic disorders. It is an in-house developed test that follows the IVDR regulations and the ISO15189-2022 standard. In such a context, any modification that could impact on the analytical performance, quality, reliability, and integrity of the results must be validated and/or revalidated.

Objective: The objective was to evaluate the impact of changes made in the implementation of the BabyDetect test, notably the relocation of testing activities, the change in software for variant analysis and interpretation, and the implementation of the NGS-WES technology.

Method: We conducted a risk analysis of the relocation and post-relocation phases, including risks related to activities associated with both the NGS-panel method and the NGS-WES method. This risk analysis led to validation plans. A global validation of the analytical workflow was performed by comparing results obtained after the relocation with those obtained before. The validation of the Qiagen-Franklin variant interpretation software was carried out by comparing results obtained with the Agilent-Alissa platform against those obtained with the new one. These validation plans were translated into actions that yielded results.

Results: This work demonstrated that neither the relocation nor the change in variant interpretation software had a negative impact on the analytical performance of the BabyDetect NGS-panel test, nor on the bioinformatics performance. Furthermore, it offers prospects for expanding neonatal screening through the introduction of the BabyDetect NGS-WES test.

Conclusion: The changes that occurred in the implementation of the BabyDetect test did not compromise the quality of the results. Therefore, the BabyDetect test can continue to be offered following the relocation of its associated testing activities. This work highlights the importance of change management to demonstrate the maintenance of compliance for our in-house developed tests throughout his life cycle.

We used Gemini AI to assist with the writing of this master's thesis, and Chat GPT to optimize our code for creating graphs in RStudio.

RESUME

Contexte : Le présent travail décrit un test de dépistage génomique néonatal de désordres génétiques pédiatriques traitables. Il s'agit d'un test développé en interne qui suit la réglementation IVDR et la norme ISO15189-2022. Dans un tel contexte toute modification pouvant avoir un impact sur la performance de l'analyse, la qualité, la fiabilité et l'intégrité des résultats doit être validée et/ou revalidée.

Objectif : Il était question dans ce travail, d'évaluer l'impact des changements qu'il y a eu dans la réalisation du test BabyDetect notamment, la relocalisation des activités d'analyse, le changement de logiciel d'analyse et d'interprétation des variants et l'implémentation de la thechnologie NGS-WES.

Méthode : Nous avons procédé à une analyse de risque du déménagement, du post déménagement, incluant aussi bien les risques liés aux activités en rapport avec la méthode NGS-panel, que ceux liés aux activités en rapport avec la méthode NGS-WES. Cette analyse de risque a débouché sur des plans de validation. Une validation du flux analytique a été effectuée de manière globale par comparaison de résultats obtenus après déménagement par rapport à ceux obtenus avant déménagement. La validation du logiciel d'analyse et d'interprétation de variants Franklin-Qiagen s'est faite par comparaison de résultats obtenus avec l'ancienne plateforme Alissa-Aligent par rapport à ceux obtenus avec la nouvelle. Ces plans de validation se sont traduits en actions ayant produit des résultats.

Résultats : Ce travail a permis de démontrer que aussi bien le déménagement que le changement de logiciel d'interprétation des variants n'ont eu d'impact négatif ni sur les performances analytiques du Test BabyDetect NGS-panel, ni sur les performances bioinformatique. Il offre par ailleurs de perspectives dans l'élargissement du dépistage néonatal par introduction du Test BabyDetect NGS-WES.

Conclusion : Les changements survenus dans la réalisation du test BabyDetect n'ont pas compromis la qualité les résultats. Le test BabyDetect peut de ce fait continuer à être proposé après déménagement des activités d'analyse liées. Le présent travail souligne l'importance de la gestion des changements dans l'optique d'une démonstration factuelle du maintien de la conformité du Test développés in House tout au long de son cycle de vie.

TABLE OF CONTENT

I.INTRODUCTION

1.Institutional and regulatory framework

This work was carried out at the University Hospital Center of Liège, a legal entity of the Unilab-Lg laboratory, as part of a quality assurance internship within the genetic biochemistry laboratory.

Unilab-Lg is a university hospital laboratory that brings together clinical biology activities, human genetics, pathological anatomy and cytology, and forensic toxicology as shown in figure 1. It has a quality management system (QMS) in place that ensures the competence and reliability of its analyses, in accordance with the requirements of ISO 15189-2022 (BELAC file no. 128 MED) and ISO 17025-2017 (BELAC file no. 128 TEST for forensic toxicology), as well as applicable European or Belgian regulatory frameworks such as Practical guideline for the implementation of QMS in accredited clinical biology laboratories version 4-2025[1], AR.1987.12.14 Royal Decree setting the standards that human genetics centers must meet (amended by the RD of January 25, 1989)[2].

The laboratory holds an accreditation issued by BELAC for several tests and NIHDI approval in clinical biology, human genetics, and anatomical pathology. Its activities are supervised by the relevant national authorities, including the FAMHP[3] responsible to ckeck the application of IVDR and Sciensano responsible to check the application of the practical directive.

Within this laboratory, the Biochemical Genetics Unit, supervised by François Boemer, is responsible for newborn screening activities and the BabyDetect genomic newborn screening test. This unit has Laboratory Developed Tests (LDTs), which are subject to the requirements of the IVDR Regulation.

In the context described above, the implementation and evolution of these tests require continuous monitoring and evaluation of their performance and regulatory compliance, particularly through validation and revalidation processes. To this end, quality assurance is involved in each of these steps to ensure the traceability of all changes, the existence of evidence maintaining analytical performance following modifications, and the alignment of changes with normative and regulatory requirements.

2.Context of Newborn Screening

Newborn screening (NBS)[4] constitutes a major public health intervention aimed at the early detection of diseases, often of genetic origin, to implement therapeutic management before the onset of symptoms. The goal is to prevent, or at the very least limit, the serious and irreversible consequences of these pathologies.

Historically, NBS was developed following the discovery of phenylketonuria[5] by Ivar Asbjørn Følling in 1934, and the development of the Guthrie test by Robert Guthrie in the 1960s. The

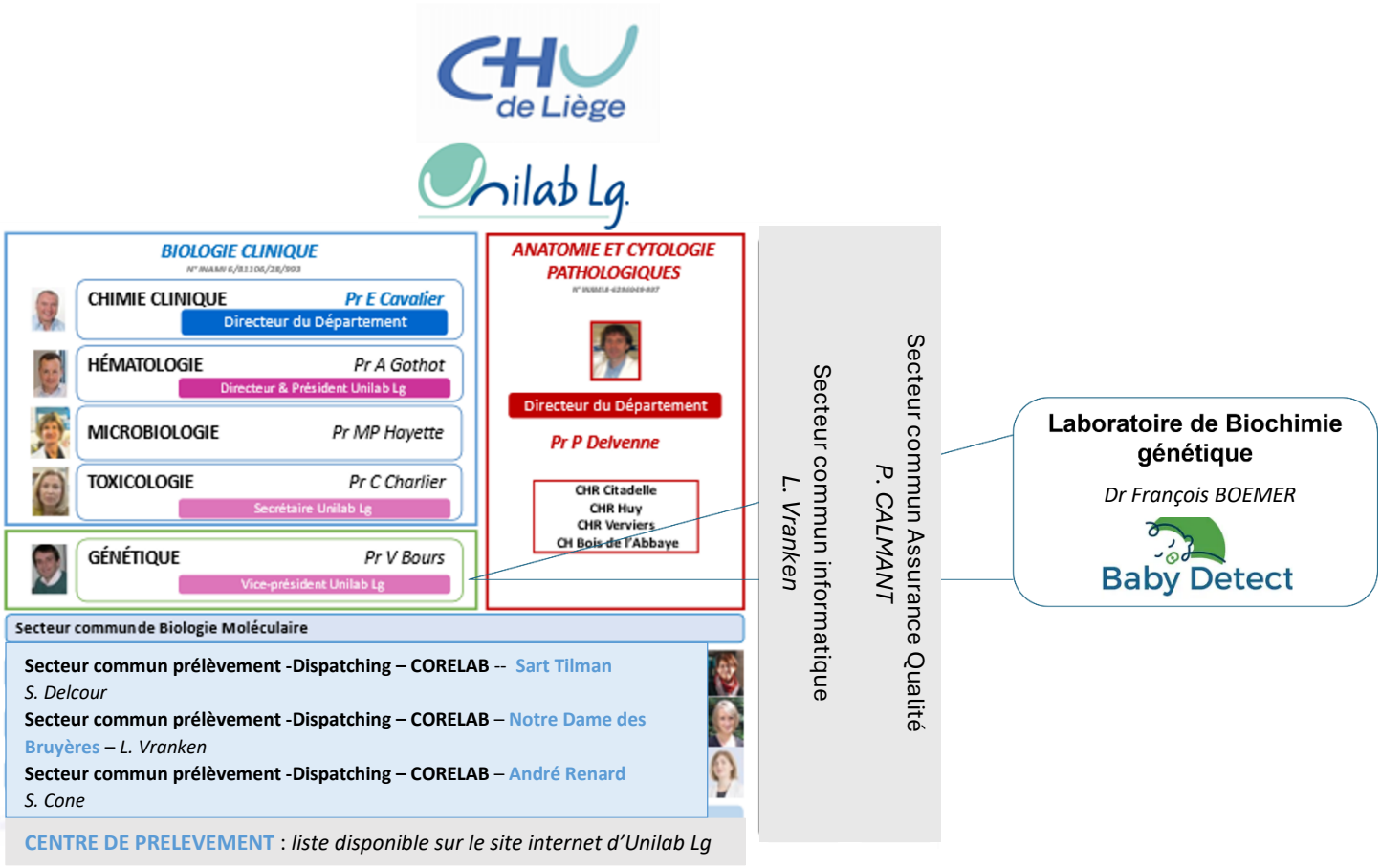


Figure 1: Organizational structure of the Unilab-Lg laboratory within the Liège University Hospital (CHU de Liège): overview of the Unilab-Lg management team, the different laboratory sectors and their respective heads, with a focus on the Department of Biochemical Genetics within the Genetics sector (From Quality Manual v25 – Unilab QMS Vivaldi).

targeted biomarker was either the urinary phenylpyruvic acid or the phenylalanine in the blood[6]. The test is based on bacterial inhibition. It uses a strain of *Bacillus subtilis*, a culture medium containing a toxic analogue of phenylalanine, and a drop of blood dried on filter paper. If the newborn's blood contains high levels of phenylalanine, it neutralizes the inhibitory effect of the analogue, allowing the bacteria to grow around the filter paper disc. These advances enabled the introduction of systematic screening using dried blood spots (DBS) on filter paper.

Biochemical assays, including LCMS, ELISA, qPCR were used until, over time, the number of screened diseases evolved using the same DBS matrix. The expansion of detectable diseases continued up to approximately 50 conditions with the introduction of tandem[7] mass spectrometry (late 20th and early 21st century). At that point, it became necessary to address the question of which criteria should be used to include a disease in the screening program.

This is how Wilson and Junger[8] defined the principles upon which screening programs still rely today. They established the selection criteria for diseases to be screened, notably their severity, the existence of a treatment, and the availability of a reliable and acceptable test. At this stage, it is necessary to distinguish between diagnosis and screening.
Indeed, screening is intended for individuals presumed to be healthy and requires further testing in the event of a suspected disease. It therefore necessitates additional investigations. In contrast, according to the Littré dictionary, diagnosis is the art of recognizing diseases by their symptoms and distinguishing them from one another. It thus consists of a clinical approach aimed at confirming or ruling out a suspected pathology presenting biological and/or phenotypic manifestations. This distinction implies specific analytical requirements, including high sensitivity, specificity, and precision, to limit false positive (FP) and false negative (FN) findings.

3.Evolution of newborn screening in Belgium – Wallonia-Brussels federation

In Belgium, newborn screening was formalized in 1968[9] and is organized under the responsibility of public health authorities, notably the *Office de la Naissance et de l'Enfance* (ONE)[10] in the Wallonia-Brussels Federation, with the support of a steering committee. Its evolution is shown in Figure 2.

Initially based on biochemical approaches, NBS now relies primarily on tandem mass spectrometry, which allows for the detection of numerous metabolic diseases through acylcarnitine profiles[11]. Following ionization (all acylcarnitine molecules produce an 85 Da ion, making it possible to identify and quantify all molecules within the same family) and the fragmentation of these molecules, they are separated according to their mass and charge. Previously, only diseases with a specific biomarker could be subject to newborn screening.

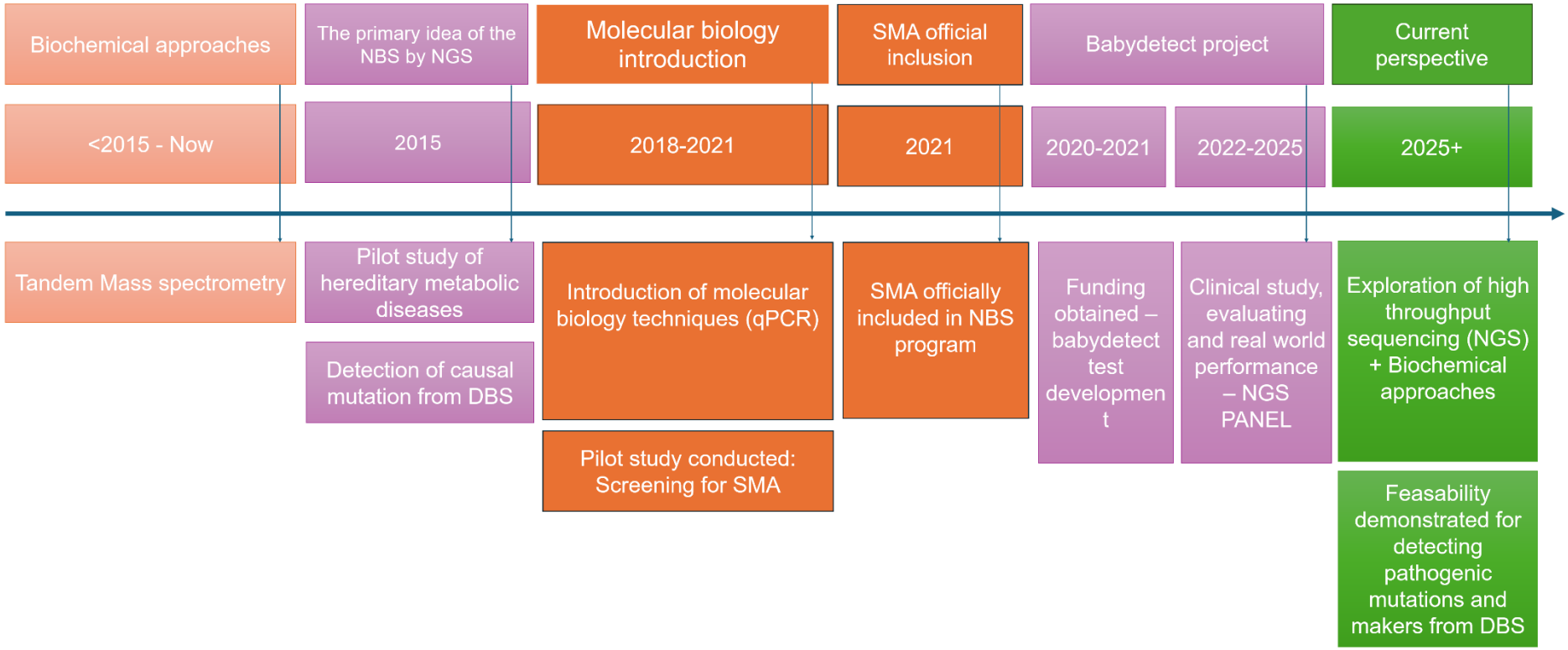


Figure 2 – Evolution of newborn screening in Wallonia-Brussels Federation (Belgium, 1968–2025). From biochemical methods (tandem mass spectrometry) to molecular biology (qPCR SMA) and next-generation sequencing (BabyDetect Test)

Next, in 2015, the concept of a genetic screening test was proposed under the title "Implementation of next-generation sequencing (NGS) technologies in the newborn screening of inborn errors of metabolism". Since then, research has been conducted, resulting in a pilot[12] study for the screening of hereditary metabolic diseases through the detection of causal mutations from DBS on filter paper. This study demonstrated that it is entirely possible to have confidence in NGS results for identifying pathogenic mutations using genetic material extracted from DBS.

This period was followed by the introduction of molecular biology techniques in the NBS. Indeed, with the emergence of new treatments for spinal[13] muscular atrophy (SMA), molecular biology technologies enabled the screening of SMA via qPCR, in a single test. Accordingly, a pilot[14] study took place between 2018 and 2021 in Liege, after which SMA was officially included (March 2021) in the newborn screening program of the Wallonia-Brussels Federation.

The idea of extending genetic screening to all treatable pediatric genetic diseases then took root, and the BabyDetect genomic NBS study was initiated, allowing the development of the BabyDetect[15] test starting in 2020, followed by a clinical study: screening of newborns between September 2022 and May 2025 to assess the feasibility, acceptability and clinical utility in a real-world screening context.

In this dynamic, the integration of high-throughput sequencing technologies has been explored[16], demonstrating the feasibility of identifying pathogenic mutations from DBS.

4.Limitations of conventional approaches

Biochemical methods, although effective, present several limitations:

- The dependence on the presence of a detectable biomarker.
- The variability of metabolites according to physiological conditions: for example, in cases of neonatal stress[17] or depending on the infant's diet.
- FP and FN: since multiple markers can be used for the same pathology, the choice of markers to consider and the most suitable technique[18] are often causes of false positives[19,20].
- The inability to detect certain genetic pathologies: for patients with more than one pathogenic mutation, biochemistry struggles to identify them all at once.

These limitations have led to the development of methods based on direct DNA analysis, allowing for the expansion of the spectrum of detectable diseases. However, DNA-based methods also have limitations; hence the importance of using these methods in combination[18,21] with biochemical methods

5.Some genetics reminders[22]

Every human being is made up of chromosomes, which carry genetic information through DNA. This DNA consists of numerous genes (the basic unit of heredity, consisting of a segment of DNA arranged in a linear manner along a chromosome), each represented by a sequence whose building block is the base. As such, DNA is a macromolecule carrying genetic information, composed of linked deoxyribonucleotides. This DNA molecule contains fundamental constituents, namely purine and pyrimidine bases, which play a key role in genetic structure and energy transfer. Purines (adenine = A and guanine = G) have a double-ring structure, whereas pyrimidines have a single-ring structure (cytosine = C, thymine = T, and uracil = U). The latter pair with purines form the DNA double helix through base complementarity. DNA can be single-stranded or double-stranded, with the strands held together by base pairing: A = T and C ≡ G. A, T, C, and G are the bases found in DNA

This DNA consists of coding regions, the exome, and non-coding regions, such as introns, regulatory elements, microRNAs, repetitive sequences, and long non-coding RNAs. The coding regions, which account for 1.5% of the genome, represent the parts of the genome that are transcribed into messenger RNA and subsequently translated into proteins.

For various reasons, DNA can undergo modifications; these are referred to as mutations or variants, which may or may not be pathogenic.

A pathogenic variant is an alteration in a gene, differing from the reference sequence, associated with an abnormal phenotype or an increased risk of developing a disease.

Among the types of pathogenic variants, we will focus on three categories:

- SNPs (Single Nucleotide Polymorphisms): These are primarily substitutions of 1bp: one type of nucleotide is present on an allele of a chromosome, while at the equivalent position on the other chromosome, there is a different nucleotide. SNPs can be categorized as follows: on one hand, there are transitions (a point mutation that transforms one purine nucleotide into another purine (A-G) or one pyrimidine nucleotide into another pyrimidine (C-T)); and on the other hand, transversions : where there is a purine on one allele and a pyrimidine at the same position on the other chromosome (A-T; A-C; G-T; G-C).
- CNVs (Copy Number Variants): These can involve large deletions (a large segment of a chromosome that has disappeared from its homologue) or segmental duplications (the presence of one or more extra copies of a DNA segment) with size from 1kb up to many Mb.

- INDELs (Insertions-Deletions): These involve the addition or removal of a portion of a DNA sequence within a gene or another region of the DNA. This may involve a single base within a gene or one or more genes (size 1 to 10 000 bp). In the context of this work, the insertions-deletions considered are those with a size less than or equal to 15 base pairs (bp).

A functional variant: these can be coding or regulatory, meaning that a genetic mutation in these 'switches' will have a direct transcriptional effect, modifying the expression level or transcript concentration of the regular gene/neutral variants.

Germline vs. Somatic mutations: These constitutive germline mutations/variants must be distinguished from somatic variants, which arise in our soma (and are not inherited from our parents) but appear through mutations during our development. An individual is mosaic for somatic mutations (which are not present in all the cells of the body, but only in the pool of cells derived from the single cell that underwent the mutation).

Exome: it is the part of the genome (DNA in all chromosomes and in mitochondria) that includes all coding nuclear DNA sequences. The human exome comprises approximately 180,000 exons that are transcribed into mature RNA. Exon is a coding sequence of DNA present in mature messenger RNA.

Hemizygous: refers to a gene normally present in only a single copy; usually an X-linked gene in a male.

6.The technological revolution: sequencing techniques

Sequencing is a method that allows DNA molecule sequences to be rendered in a readable format, enabling the exploitation of the information they contain. Several technologies are distinguished based on their scope: gene panels, Whole Exome Sequencing (WES), or Whole Genome Sequencing (WGS). There are four generations[23] of sequencing: the first being Sanger sequencing – automated between 1980 and 1990; the second being sequencing-by-synthesis via fluorescence or pH measurement, which includes technologies such as Illumina and Roche since 2005; the third generation involves single-molecule sequencing – introduced between 2010 and 2012; and the fourth generation is performed using nanopore technology – the emergent technology since 2020.

- **Sanger Sequencing[23]:** This method relies on the incorporation of dideoxynucleosides (ddNTPs), known as terminators, during DNA synthesis. These molecules, which lack a hydroxyl group at the 3’ position, interrupt the elongation of the strand currently being synthesized. The reaction is performed in the presence of deoxynucleotides (dNTPs) and specific ddNTPs (A, T, C, or G), thereby generating a population of DNA fragments

of varying sizes, each ending with a specific nucleotide. These fragments are then separated by electrophoresis according to their size, allowing the DNA sequence to be reconstructed by reading the fragments in order.

- **Sequencing-by-synthesis**[23] occurs in four stages: it begins with preparation, where the samples are fixed on a flow cell coated with known nucleotides.

The single-stranded DNA fragments to be sequenced have adapters complementary to these nucleotides. This is followed by bridge PCR: the DNA polymerase synthesizes a complementary strand from primers fixed to the flow cell. When heated, the non-fixed strand detaches while the other strand remains attached. This remaining strand binds to another oligonucleotide on the flow cell, forming a 'bridge.' Upon cooling and adding polymerase, synthesis of a new strand occurs again. By repeating these operations, clusters of identical DNAs are formed. Next, sequencing begins by adding four reversible terminator nucleotides (A, C, G, T), each labeled with a different colored fluorophore. A terminator blocks polymerization (due to the lack of a 3’ OH group). All fluorophores are excited, and the color emitted by each cluster is recorded with a camera, thereby identifying the base. Through a chemical wash, the fluorophore is removed and the OH group is restored to allow the addition of the next nucleotide. The cycle is repeated (wash-and-scan) to read the sequence base by base.

6.1 NGS Panel and NGS-WES

Next-Generation Sequencing (NGS) uses the second-generation sequencing technology. It primarily enables the simultaneous analysis of multiple genes for gene panels[22], the genome, the exome, and the transcriptome, thereby offering an increased capacity for detecting genetic diseases.

When referring to gene panel sequencing, we are describing simultaneous molecular tests on different genes that share the same clinical phenotype, a similar phenotype, or different phenotypes. The genes included in the panel vary depending on the area of interest[24], the clinical target[25], and its accessibility. WES is a similar technique with the difference that specific exons are selected in the case of a filtered exome, or the entire exome is sequenced. Initially used in a diagnostic context, NGS has been progressively introduced into screening strategies, particularly within the framework of pilot[12,26,27] projects such as BabyDetect.

Both Panel and exomes needed bioinformatic software to treat data and software to analyze and interpret them. To categorized the variants, The American College of Medical Genetics and Genomics (ACMG), in collaboration with the Association for Molecular Pathology (AMP)

and the American College of Pathologists, proposed a classification of variants into five groups[28,29]: pathogenic, likely pathogenic, benign, likely benign, and of uncertain significance. They also proposed an interpretation of variants for mendelian disorders (caused by mutations in a single gene) in a diagnostic context; they note that this interpretation is not always adequate in the context of healthy patients. Despite the advances achieved through NGS, this technology also has its limitations.

6.2 Challenges and Limitations of NGS in Screening

Despite its advantages, NGS presents several technical and conceptual limitations:

- On an analytical level: incomplete coverage of certain regions or the total lack of coverage in some areas. Genetic heterogeneity is a limitation because, for the same disease (e.g., deafness or cardiomyopathy), many different genes can be involved; consequently, targeted panels may miss them. There are also the risks of sequencing errors and of false positives[17] in cases where a heterozygous mutation is found in a cis position in the patient (limitation of short-reads sequencing).
- On a bioinformatic level: the dependence on analysis pipelines and the management and storage of massive amounts of data.
- On a clinical level: Interpretation of sequenced data of asymptomatic newborn population, no phenotypic information is available. the difficulty of variant interpretation and the management of variants of uncertain significance[30] in the absence of a consensus.
- On an ethical level: the management of incidental findings[31]: In a diagnostic setting, there is a list of genes for which it is recommended to search for and report pathogenic variants, even if they are unrelated to the initial reason for the test. Integrating this concept when reporting pathogenic cases within a screening framework is not always straightforward.

The BabyDetect study illustrated these challenges, notably through the identification of false positives and false negatives, requiring an in-depth analysis of the causes and limitations of the method.

NGS constitutes a significant paradigm shift. Unlike biochemical methods, it involves several successive complex stages, from sample preparation to bioinformatic analysis on many samples simultaneously. This complexity necessitates the implementation of a management system to maintain the quality of results, to control sources of variability and guarantee the reliability of the findings. To this end, the quality team, in collaboration with scientists and

technicians, ensures that the various stages are carried out in compliance with current regulations, while maintaining thorough documentation and traceability of all operations.

7.Validation and Quality Assurance of NGS Methods

In a laboratory accredited according to ISO 15189, every analytical method must be validated before its implementation into routine use. Validation consists of demonstrating, based on objective evidence, that the method meets the defined performance requirements.

7.1 Validation Requirements According to ISO 15189: 2022

ISO 15189[32] is an international standard applied to medical laboratories, addressing the requirements and competencies related to these facilities. It is also utilized by regulatory authorities and accreditation bodies to grant laboratories recognition of their competence in the services they provide.

Under the ISO 15189:2022 standard, the validation of analytical methods constitutes a fundamental requirement, ensuring the reliability of the results provided to patients. It is part of a quality management system (QMS) that covers the entirety of the pre-analytical, analytical, and post-analytical processes.

In the case of NGS testing, validation is multidimensional and comprises analytical validation, bioinformatic validation, and clinical validation. These different dimensions of validation are part of a global quality management approach that integrates risk management, traceability, reproducibility, as well as internal and external quality controls. It is essential to determine when validation is required according to ISO 15189:2022.

7.1.1Circumstances for the Validation or Revalidation of an Analytical Method

According to ISO 15189:2022, in section 7.3.3, any method used outside of its originally intended scope, any laboratory-developed method, or any major[33] modification to an initially validated method must undergo validation or revalidation. In view of this requirement, since the BabyDetect test is an LDT method on one hand, and has undergone significant modifications on the other, it requires both validation and revalidation.

7.1.2 Requirements for the Validation/Revalidation of Analytical Methods

The validation of a method requires competent personnel and the appropriate skills to perform the validation. This requirement is particularly significant in the context of technology such as NGS, which is quite complex at the technical (operational/handling), scientific, and clinical levels. Therefore, Quality Assurance (QA) ensures compliance with this requirement during the validation process and verifies the existence of objective evidence demonstrating the performance of both scientists and technicians. This is achieved, for example, through the

establishment and recording of competency files, as well as the tracking and maintenance of staff skills.

Furthermore, for the traceability, the standard requires all records related to the validation or revalidation of methods to be stored. This specifically includes the dossier of a validation file containing the validation procedure used, the scope of the method's application, the determination of the method's performance specifications, the results obtained by the method, and a conclusion regarding the suitability of the method for its intended use.

QA also intervenes at this stage, notably by participating in the development of a risk analysis. This serves as the starting point for establishing the best strategy to validate or revalidate methods, while keeping in mind the identified risks and their associated control measures. Furthermore, QA participates in the creation of the validation file, ensuring its overall compliance.

In the context of the BabyDetect test, in addition to being subject to the requirements of the ISO 15189:2022 standard, it is also an in vitro NBS device developed by the laboratory (LDT) under the IVDR.

7.2 Laboratory-Developed Tests – related requirements

The IVDR is a European regulation that sets high standards for quality as well as safety for in vitro diagnostic medical devices, to address common safety challenges related to these devices. It ensures that data derived from device performance studies are reliable and robust, and that patient safety is preserved.

According to ISO 15189:2022, an in vitro diagnostic medical device is a device, used alone or in combination, intended by the manufacturer for the *in vitro* examination of specimens derived from the human body, solely or principally to provide information for diagnostic, monitoring, or compatibility purposes. This includes reagents, calibrators, control materials, specimen receptacles, software, and related instruments, apparatus, or other articles. The IVDR[34] further adds that any test providing information on predisposition to a medical condition or a disease, such as genetic testing, is an in vitro diagnostic medical device. Consequently, the BabyDetect test can be considered an IVD. However, the BabyDetect test is an LDT and does not adhere to all the constraints of the IVDR. The ISO/DIS 5649:2023 standard, among others, addresses issues related to LDTs.

As such, ISO/DIS 5649 is an international standard applied to medical laboratories regarding the concepts and specifications for the design, development, implementation, and use of LDTs. It ensures the quality, safety, performance, and documentation of LDTs based on their intended use for the diagnosis, monitoring, prevention, or treatment of medical conditions.

7.2.1 Definition and Lifecycle of an LDT

According to ISO/DIS 5649:2023, an LDT is a test developed or modified (from a commercial solution) and used within laboratories on samples of blood, biological fluids, or tissues, as well as on specimens derived from the human body, for a specific objective that existing solutions fail to achieve.

In the framework of an LDT, compliance with their requirements mandates a continuous demonstration of performance and a justification for their use throughout their entire lifecycle (see Figure 3). This lifecycle consists of:

- A feasibility study phase: this involves brainstorming and evaluating a potential solution to an identified problem. In the case of BabyDetect, the goal was to expand the range of detectable diseases by utilizing a technology that does not depend on a specific biological marker.
- A design and development phase: this addresses the question of how the device will function to achieve its objective, which in this case, the screening of genetic diseases in newborns using DNA, without the need for traditional markers.
- A pilot study phase via performance evaluation: this consists of testing the device that has been developed. In the case of BabyDetect[12], this phase took place between 2020 and 2025.
- A validation phase: this confirms that the developed device is indeed fit for purpose and provides satisfactory results in accordance with its intended use. This phase has been completed for the NGS-panel (validation file: BGE.BABYDETECT.ANA.A01 v2 – Vivaldi QMS) and is the focus of this work for the NGS-WES.
- An implementation and monitoring phase: following validation, the test is used in routine practice and requires monitoring to continuously verify its compliance. As daily monitoring for the BabyDedect targeted NGS test, we have internal quality control at each run and key performance indicators. During this phase if there is a change like in this case, the moving or the change of the variant's interpretation software, the BabyDetect test must be revalidated after a risk analysis performed. In this case we go back to the previous phase.

  Beyond the technical aspect, monitoring also encompasses regulatory observation. The device must remain aligned with current regulations; consequently, the QA team, in collaboration with the technical team, must stay up to date regarding various normative or regulatory changes throughout the device's lifecycle. It is therefore a matter of strengthening the regulatory monitoring by regularly consulting the FAMHP[35] website and the European Commission's pages containing the MDCG[36] guidelines on

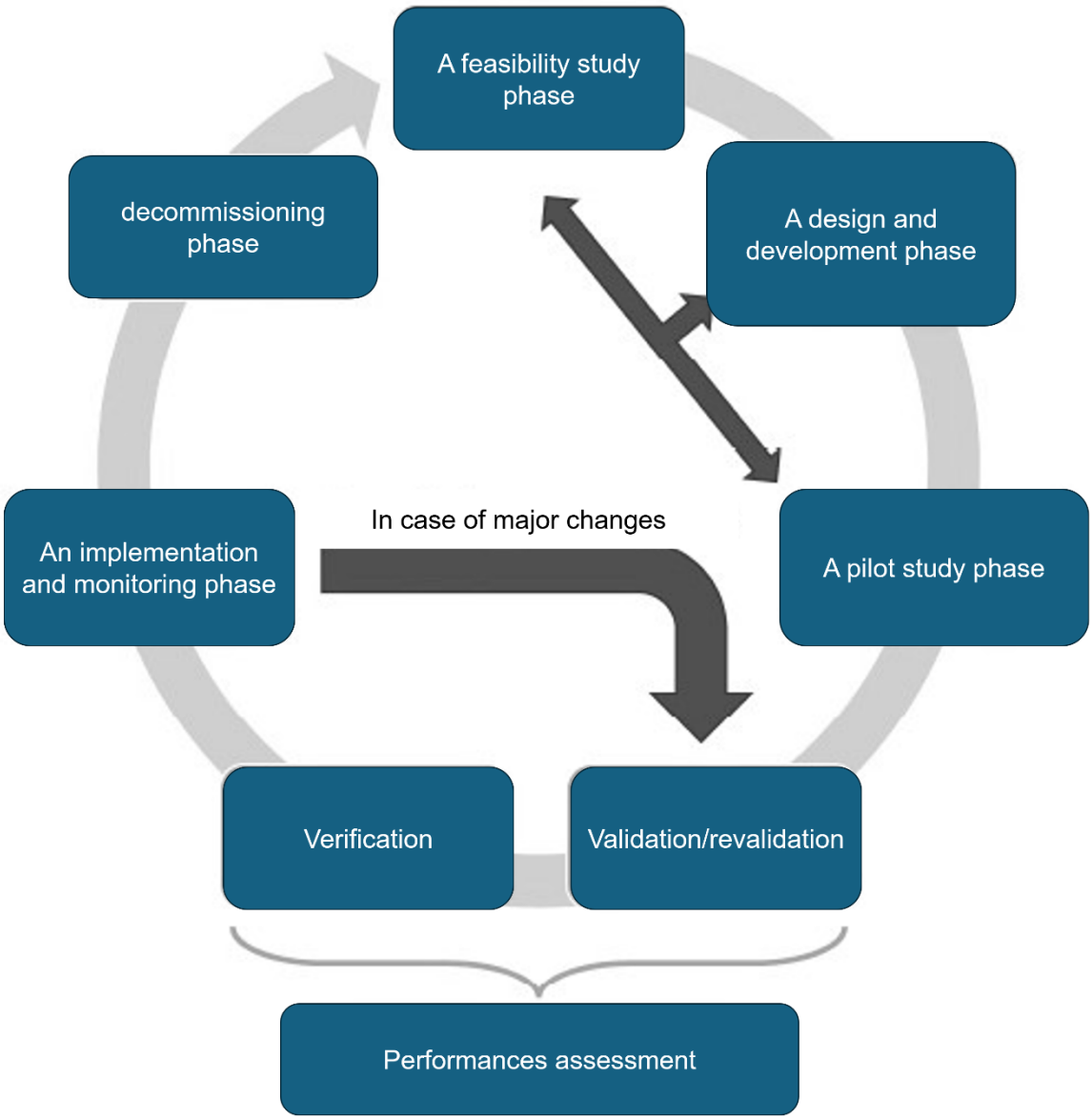


Figure 3: LDT life cycle: the sequential phases of an LDT, from initial motivations to withdrawal (Adapted from ISO/DIS 5649:2023).

medical devices and IVDs, which are essential for following updates to the IVDR and MDR.

- And finally, a decommissioning phase, if applicable: A test that no longer meets the criteria initially defined when it was established and put into routine use in the laboratory must be withdrawn from routine use. The withdrawal can be for several reasons: notably the introduction and use of a new LDT or a new method making the previous one obsolete, or the introduction of new technologies.

Regarding the BabyDetect test, the various stages proceeded through validation and implementation for the targeted NGS-panel workflow. Following a modification during the implementation stage, it is necessary to perform a revalidation, after which the workflow will continue its operation. As for the NGS-WES implementation, the BabyDetect test has also followed every stage up to the focus of this work—namely, the method validation stage.

7.2.2 Recognition Criteria and General Requirements for LDTs

Recognition Criteria for an LDT

- It is entirely designed by the laboratory, and/or
- It utilizes a commercial test but with significant modification to the manufacturer's intended use, or lacks full coverage by an IVDR certification, and/or
- It involves a test initially developed for Research Use Only (RUO) but validated for human diagnostic use, and/or
- There is an absence of IVDR certification for an essential component (reagent, instrument, software), except for specific exemptions.

General Requirements: to avoid being fully subject to the IVDR regulation, an LDT device must comply with the following points:

- It must be used solely in-house, meaning it cannot be transferred from one entity to another. In other words, under its current status, the BabyDetect test could not be performed outside the CHU (University Hospital). This rule has been in application since 26 May[48] 2022.
- A well-established Quality Management System (QMS) must be in place to oversee the activities related to the performance of the test. This quality system must incorporate compliance with the ISO 15189 standard or any national equivalent.
- Provide evidence that no commercial solution exists that can achieve the results obtained by the proposed method, or that available commercial solutions do not address the specific problems that the proposed method solve.This will apply from 31 December[48] 2030.

- To be able to provide information regarding the justification for the manufacturing of the device (for risk class C and D in Belgium), any modifications made, and its intended use. Will apply from 31 December[48] 2030.
- To be able to make a public declaration identifying the manufacturer of the device and stating whether it is compliant. In application since 26 May[48] 2024.

7.3 Some definitions

Validation can be analytical[28], bioinformatic, clinical, or IT-related. Within the scope of this work, we will address analytical and bioinformatic validation. Below are a few key concepts to consider:

- An analytical method[37] corresponds to the set of explicitly described operations used during the performance of an *analysis* in accordance with a given protocol.
- An analysis[37] is the set of operations intended to determine the numerical value, the textual value, or the characteristics of a property. An analysis can be quantitative: when the value to be determined is numerical; or qualitative: when a textual value or the characteristics of a property are determined.
- EQA[37] (External Quality Assessment), also known as Proficiency Testing (PT), is an evaluation of participant performance against pre-established criteria by means of inter-laboratory comparisons (the organization, execution, and evaluation of measurements or analyses on the same or similar materials by two or more independent laboratories under predetermined conditions).
- Internal Quality Control[37] (IQC): an internal procedure that monitors the testing process to verify that the system is functioning correctly and provides confidence that the results are sufficiently reliable to be released. Quality Metric[38] is an example: a measurable data point, recorded at the input, at intermediate points, and at the output of the pipeline, used to evaluate its performance both initially and continuously, for every run and for every sample.
- GIAB: It is a reference DNA sample provided by the National Institute of Standards and Technology. In our context, HG002 (N24385)- GIAB was used. It was characterized using five different sequencing technologies within the framework of the Global Alliance for Genomics and Health consortium. This characterization includes eleven WGS analyses and three WES analyses, combined with the most advanced variant detection tools available at the time of its characterization. The reference data associated with this sample includes: a VCF file listing all the variants identified with a high level of confidence (called in the BabyDetect case GIAB Gold) and a BED file describing the

regions of the genome that could be reliably characterized. These regions are called high confidence regions (HCR).

- Pipeline[38]: An automated computational processing chain consisting of a suite of software tools a succession of programs that analyze raw data from the sequencer, assemble it, and translate it into interpretable sequences. These software tools or programs may be developed in-house or sourced externally (commercial or open source).
- The DRAGEN concept[39]:It is a bioinformatic platform designed to transform raw data after the sequencing phase, with a high performance, into genomic insights.
- Run: A single execution of a test consisting of a set of samples sequenced in parallel on the NGS sequencer.

7.4 Validation Process

Method validation is governed by a set of parameters. Each laboratory drafts and follows a procedure appropriate for the validation of its methods within the scope of its medical analysis activities. The elements presented here are based on the COFRAC[38] validation guide. Validation applies both to new methods and to major changes made to an existing method.

The various elements to be considered for the validation of an analytical method according to the literature are:

- Process description: this includes general data regarding the analysis, its associated procedures, a summary of the software used, and their respective version numbers.
- Description of the method to be validated: Consists of providing the validation objectives; The principle of the method to be validated and its objective; The choice of the method/the change made to an old one and the matrix on which the validation is performed.
- Implementation: this covers the operators who performed the validation, the operating procedure and the reference to its associated protocol, the study period for the method, the initial date of routine implementation, and the name of the supervisor who authorized the commissioning.
- Risk management: when the method involves specific risks that do not appear in the global risk analysis, this section must demonstrate that any risks associated with this method are effectively managed.
- Identification of performance criteria to be evaluated: this follows the risk analysis.
- Evaluation of analytical performance: it describes the nature of the sample, the means of external quality assessment, the link to the storage location of EQA results, the validation parameters, and the validation results.

- Conclusion and report preparation

8.Study objectives

The integration of NGS into newborn screening represents a major advancement, but it is accompanied by increasing complexity in analytical, bioinformatic, and organizational processes.

Within the framework of the BabyDetect test, several modifications have been introduced:

- Firstly, the laboratory relocation: currently, the BabyDetect test is performed using an NGS panel in a dedicated rented location (GIGA building). Since the cost of leasing the premises represented an additional financial burden, it became necessary to relocate the BabyDetect analyses to the Unilab-Lg platform. Plus, the BabyDetect test is not more in study phase; It is proposed as a CHU test, so it is better for the test to be performed in the Unilab platform.
- Furthermore, the change in interpretation software: the laboratory received notification that the interpretation of variants, which was previously performed with the Alissa Interpret software (Agilent, reference), could no longer be conducted since the Alissa Interpret is being decommissioned. Consequently, interpretation activities had to be migrated to the Franklin platform (Qiagen reference).
- Last but not least, the evolution of the panel toward a filtered exome: the BabyDetect test aimed to expand by offering the possibility of screening even more diseases/ genes than it currently does. Given the flexibility in terms of adding and removing genes, among other factors, and the broader range of diseases to be screened, NGS-WES was the solution to meet this need.

We would like to point out that, until now, the BabyDetect test targeted 407 genes. Within the context of compliance with the IVDR regulations and ISO 15189 accreditation, section 5.5.1.3 states:

“If modifications are made to a validated analytical procedure, the impact of these changes must be documented and, where appropriate, a new validation must be performed.” Consequently, these modifications require a rigorous demonstration of their impact on the quality of results. Therefore, the problem statement of this work is as follows:

How can the feasibility and compliance of BabyDetect newborn screening tests, using NGS-panel and WES, be guaranteed following technical and organizational changes within an ISO 15189 accredited environment?

The primary objective of this work was to demonstrate the continued performance, reliability and quality of the BabyDetect test following the implementation of analytical, bioinformatic and organizational changes, through a validation and revalidation process conducted in accordance with ISO 15189 and IVDR requirements.

To achieve these objectives, the specific tasks were as follows:

- To conduct a risk analysis related to the laboratory relocation and NGS-WES
- To formalize a relocation plan
- To perform a revalidation of the current NGS-panel method following the relocation
- To validate the new variant analysis and interpretation software: Franklin
- To validate the NGS-WES method

9.Work Plan

To meet these objectives, I was trained in internal audit, I've been out in the field. Our work has been organized into four sections. Each section is divided into two main parts: the first focuses on the revalidation of the newborn screening method using a targeted panel, and the second focuses on the validation of the exome-based method.

- For the Materials and Methods section:
    - In the first part**:** we present the risk analysis we conducted regarding validation and revalidation; We then outline the revalidation plan that we established. Next, we present the materials and the various strategies deployed to revalidate the NGS-Panel method. Note that the Franklin software was used for interpreting variants detected both with the targeted panel and with the filtered exome.
    - In the second part: we present the materials used and the strategies implemented for the validation of NGS-WES.
- For the Results section: we present the results following the same logic: first, the results related to the revalidation, and then the results related to the second part (NGS-WES validation). We also explain the approach chosen to validate the new variant analysis and interpretation software: Franklin.
- For the Interpretation and Discussion section**:** We explain our findings and compare them with available literature data where applicable. We also explore the implications of these results for future public health deployment.
- Lastly, we conclude.

II. MATERIAL AND METHOD

1.Risk Analysis

The risk analysis was performed using a qualitative method based on the principles of the **FMEA** method (*Failure Mode and Effects Analysis*). It is an in-house method adapted to the need and the reality of the laboratory. We conducted the moving risk analysis according to GNT.ORG.GES.A39 version 1, where the risk matrix was described as follows (see 1.2). The first table in the result section for risk analysis took into consideration that previous matrix. During our work, we attended an external **BELAC** audit which identified a non-conformity (registered under number NC110148 in the Genetic Biochemistry-BabyDetect sector of the laboratory's quality management system: VIVALDI P2F) in the risk analysis procedure. This non-conformity concerned a non-patient-oriented risk analysis. We conducted the following risk analysis by integrating this auditor's remark. The modifications were proposed, accepted and published the 08 Junv2026 in Vivaldi QMS in the GNT.ORG.GES.A39 version 2 file. The recording files for the risk analyses were also modified in P2F. The post-moving table in result section takes into consideration the new matrix described below.

The risk calculation involves risk identification, to which a three-level probability of occurrence was assigned concerning BabyDetect analysis, a three-level detectability assessment, and the determination of the consequences (C) that could result from the occurrence of the risk (also graded on three levels). That being said, the risk calculation is the combination of 3 parameters: the probability of the risk (Probability), the detectability of the risk both before and after its occurrence (Detectability), and the potential impact of this risk's occurrence on the analysis, the operators, and the patient (Severity). Each hierarchization described here concerns the BabyDetect test.

1.1 Hierarchy of Levels for Each Parameter

- Probability (P)

- Level 1 = Low = low probability (e.g., observed per four-month period/semester, or even over a span of several years)

- Level 2 = Medium = medium probability (e.g., observed quarterly or within a two-month interval)

- Level 3 = High = high probability (e.g., once or more than once a month, or even weekly depending on the analysis)

- Detectability (D)

- Level 1 = High = High/good early or factual detection

- Level 2 = Medium = Moderate early or factual detection

- Level 3 = Low = Low/poor early or factual detection

- Severity (S)

- Level 1 = Low = No consequences for the operators, the analysis result, and/or the patient (e.g., delays/incidents with no impact on the TAT)

- Level 2 = Medium = Indirect consequences for the patient and/or the operator, direct consequences for the analysis with no impact on the final result (e.g., prolonged TAT; unfeasible analysis)

- Level 3 = High = Direct consequence on the operator's health, and/or the analysis result, and/or the patient's health (e.g., accident/incident during the stages of the complete pre-analytical/analytical/post-analytical process; erroneous result leading to inadequate patient management or clinical decision, leading to deterioration of the patient's condition or death)

In summary:

Criteria	**Level 1**	**Level 2**	**Level 3**
Probability	Low	Medium	High
Detectability	Low	Medium	High
Severity	High	Medium	Low

1.2 The Decision Matrix

Criticality was determined across three levels, considering the risk control methods already in place on the Unilab platform and within the general BabyDetect workflow, according to the following matrix: (GNT.ORG.GES.A39 version 2)

Criticality index (CI)	**Criteria**	**Recommended measures**
Low	• All criteria are level 1, • 1 of level 2 + 2 criteria of level 1 • 1 of level 3 + 2 criteria of level 1	Controlled risk and no immediate need for additional control measures.
Medium	• 1 level 3 criterion + 1 level 2 • 2 to 3 level 2 criteria • 2 level 2 + 1 level 1	Implementation of additional control measures within a defined timeframe (or justification if this is not possible).
High	• 2 to 3 level 3 criteria. • 1 one level 3 criterion + 2 level 2	Immediate and mandatory control measures.

Previously the matrix was as follows: (GNT.ORG.GES.A39 version 1)

Criticality index (CI)	Criteria	Recommended measures
Low	All the criteria are level 1, or only one criterion is level 2 (no level 3 criteria allowed).	Controlled risk and no immediate need for additional control measures.
Medium	Only one level 3 criterion and/or more than one level 2 criterion.	Implementation of additional control measures within a defined timeframe (or justification if this is not possible).
High	More than one level 3 criterion.	Immediate and mandatory control measures.

1.3 Risk Control and Assessment of Mitigation Measures

Once the risks were identified, we were able to highlight both the controlled risks and those requiring further measures to achieve risk reduction. For the risks requiring the implementation of new measures, we assessed the feasibility of the actions to be deployed based on three levels:

-Level 1 = Low = Relates to actions for which the necessary resources and means are readily accessible and implementable. For these actions, the overall risk could be reduced by one or two levels, depending on the current risk level.

-Level 2 = Medium = Relates to actions for which the necessary resources and means are moderately accessible and less straightforward to implement. For these actions, the overall risk could be reduced by one level.

-Level 3 = High = Relates to actions for which the necessary resources and means are difficult or impossible to implement. For these actions, alternative mitigation methods could be proposed, or a decision must be made regarding risk acceptability.

In summary:

Criteria	Level 1	Level 2	Level 3
Feasibility of actions (F)	Low	Medium	High

This approach provided us with an overview of the residual risk level that could be achieved once the measures are implemented.

1.4 Risk Monitoring

The purpose of risk monitoring is to evaluate the effectiveness of the actions implemented both to mitigate risk in situations where the criticality index remains high, and to ensure that low-level risks are maintained at their current level. This monitoring will be conducted annually, and the risk assessment will be updated accordingly based on the observations made.

2. Revalidation after relocation

2.1 The Validation Plan

Based on this risk assessment, a post-relocation validation plan was developed. It will be presented in parallel with the results associated with each stage of the revalidation process.

2.2 The NGS-PANEL Workflow – Material and Methods

We observed the BabyDedect flow from sample reception to sequencing, which allowed us to update the standard operating procedure BGE.BABYDETECT.ANA.A06, identify some risks, and update the risk analysis sheets. The workflow of the NGS panel was validated in 2022. Here we present the steps of this workflow, which were taken up again as part of the revalidation after the move. The validation file related to this validation can be viewed in the laboratory's quality management system in Vivaldi Document Control under the name BGE.BABYDETECT.ANA.A01.

2.2.1. Sample Reception and Preparation via Twist Technology

For most cases, samples were collected on filter papers (Goldcarte BabyDetect) within 48 hours postpartum. Once the samples were delivered to the Unilab central dispatching unit, they were retrieved and logged into the system by the BabyDetect secretarial staff using the laboratory information management system (LIMS) GLIMS which is presented below.

The preliminary step to DNA extraction was then performed: punching, using a puncher [insert ref]. This was followed by a sample lysis phase using the Illumina Lysis Kit (ref: 20034185) and Illumina Purification Beads (ref: 20036210). Subsequently, the lysed samples were loaded onto a deep-well plate in the Nuclocube 1 automated system (Biocubik), and the extraction process was executed.

Due to its robustness and its streamlined workflow for both exome and targeted panel library preparation, TWIST technology was selected. It also offers an optimized exome library that enables the sequencing of a maximum number of patients on a single flow cell; this same

substrate is highly efficient for generating panels that specifically target selected genes or regions. Following DNA extraction, the subsequent steps were performed:

- Enzymatic DNA Fragmentation: Performed in Pre PCR room with a Thermocycler using the Twist Library Preparation EF kit 1, 2.0 (Twist Bioscience – ref: 101058). This step serves a dual purpose: end repair and A-tailing. Fragmentation generates "sticky" DNA ends. Blunting regularizes these ends to produce "blunt" ends by adding or removing specific bases. Subsequently, A-tailing consists of the enzymatic addition of an adenine (A) base to the 3' end of the repaired fragments, which is essential for the downstream ligation step.
- Adapter Ligation: performed in pre-PCR room. Universal adapters were ligated by adding the Twist Universal Adapter System TrueSeq compatible Plate D (Twist Bioscience – ref: 101311). Adapters are short sequences containing a thymine (T) base that bind to the fragmented DNA via complementary base pairing (A-T), facilitating subsequent sequencing. This was followed by a purification step of the adapter-ligated DNA using the Twist Binding and Purification Beads kit (Twist Bioscience – ref: 100983).
- PCR Amplification: Utilizing the same kit mentioned in the previous step, a PCR amplification stage was performed in post PCR room using primer pairs containing specific indexes. These indices uniquely identify each DNA sample, and therefore each individual specimen. A purification step of the amplified DNA followed. At this stage, the products are referred to as DNA libraries. Each library was assayed and quantified using the Qubit and the Qubit dsDNA Broad Range Assay (Invitrogen – ref: Q32853).
- Sample Pooling: To reduce the number of tubes handled and minimize manual handling errors, multiplex pooling was performed using purification beads, combining 8 samples per pool to yield a total of 6 pools. The objective was to achieve an approximately uniform library mass across all patients (187.5ng/patient).
- Hybridization and Capture: Conducted using the Twist Hybridization and Wash kit (Twist Bioscience – ref: 101025), Twist Custom Panel (Twist Bioscience – ref: 101001), and Twist Universal Blockers (Twist Bioscience – ref: 100578). This step involves using panel probes conjugated with biotin molecules. These probes target the regions of interest and are introduced alongside blockers to prevent non-specific universal adapter-to-probe hybridization, thereby promoting target DNA-to-probe hybridization.
- Washing and Enrichment: Upon completion of the hybridization, washing was performed using Streptavidin beads to optimize probe-DNA complex binding to the

beads, thereby optimizing purification efficiency. It is noted that the probes incorporate biotin molecules, which exhibit an exceptionally high affinity for streptavidin. Post-capture PCR amplification was then carried out using a master mix containing KAPA HiFi HotStart ReadyMix (ROCHE / Kapa Biosystems – ref: 7958927001). This PCR amplification step allowed for the enrichment of the targeted regions. Following this step, the library pools were quantified using the Qubit and the Qubit dsDNA HS Assay (Invitrogen – ref: Q32854).

2.2.2 Sequencing and secondary bioinformatics analysis

To prepare the sample for loading onto the flow cell, a pool combining all libraries into one was created. The DNA was denatured using NaOH, and then a spike-in control (PhiX) was added to further diversify the sequences and improve the overall run quality. The pools were diluted to achieve equitable library concentration, homogeneous loading, and a more precise calculation of the concentration deposited onto the flow cell.

The revalidation was performed on a 48-sample plate using NextSeq150 v2.5 HM (74-10-10-74) flow cell and The NextSeq 500/550 High Output Reagent Cartridge v2; on the Illumina NextSeq 550 sequencer. The sequencing can also be performed on a Novaseq.

2.2.2.1 Bioinformatics Pipeline: secondary bioinformatic analysis

The bioinformatics pipeline used was Humanomics, an in-house raw sequencing data analysis program developed by the genetics laboratory's bioinformatics team. This pipeline had been previously validated. The pipeline steps are described below. We focused on SNPs and indels of up to 15 bp. The pipeline steps are as follows:

- Raw sequencing data retrieval: Data generated by the sequencer was retrieved and processed through a demultiplexing tool to obtain data in FASTQ format, containing all reads with quality scores. This data underwent a global quality control assessment to evaluate run quality.
- Alignment: Once quality scores were satisfactory, these FASTQ files were input into Humanomics to generate VCF files. Each short read was mapped to its position using the reference genome (hg19 in this case). This step produced BAM files containing the aligned reads using the bwa-mem aligner.
- Sorting, indexing, and variant calling: Using a dedicated tool, the software scans each position of the genome in the cleaned BAM file and determines, based on statistical models, whether a difference from the reference is a true biological variant or merely a sequencing error. The software assigns a probability score (e.g., QUAL in a VCF file). If this score exceeds a predefined frequency threshold, the variant proceeds to the next

step (it is called — i.e., if something differs from the normal sequence, it is retained for later determination of pathogenicity). After this step, VCF files were obtained and subsequently input into the analysis and interpretation software.

2.2.2.2 Quality Parameters

To proceed to the next step, quality controls were performed based on a set of parameters. For each parameter, the acceptability thresholds initially defined during the primary validation were used for the post-relocation validation. Criteria were defined based on literature data and simulation experiment data, and on the distribution analysis of the results obtained during the three initial validation runs.

Outliers are identified because they show an excessive relative deviation from quartiles. These values are identified if they deviate by more than 1.5 X IQR from the first and third quartiles, meaning if they are lower than Q1 - 1.5 X IQR or higher than Q3 + 1.5 X IQR. Depending on the desired tolerance level, the 1.5 X factor can be doubled.

The parameters monitored and their acceptability thresholds were as follows: The parameters in italics and bold are VIP metrics. For each of the parameters, when a value is outside the acceptable thresholds, other parameters are investigated in collaboration with the bioinformatics team. If there's no significant impact, the analysis continues; otherwise, the analysis is redone for the sample whose parameters are critical.

- For Library Preparation and Capture

-MEAN_TARGET_COVERAGE: This is the average number of times each base in the target region was sequenced. It reflects the reliability of variant calling. For this parameter, the lower limit is 117X.

-SELECTED_BASES_PCT: This is the percentage of aligned bases located on or near the targeted regions during capture. This parameter verifies that the capture step was successful. The lower limit is 78%.

-MEAN_INSERT_SIZE: This is the average size of the fragments prepared after the DNA extraction step. It corresponds to the DNA size between two reads when sequencing in paired-end mode. A size that is too small or too large indicates a problem during fragmentation (e.g., poorly sheared DNA or initially degraded DNA). The lower limit for this parameter is 152 bp and the upper limit is 266 bp.

-FOLD_80_BASE_PENALTY: This parameter determines whether the coverage of the targeted regions is uniform. If not, it indicates by how much the mean coverage must be

multiplied by 80% of the bases to reach the mean coverage of the targeted regions. The closer the value is to 1, the better. The maximum value is 3.9.

-TARGET_BASES_30X_PCT: *The percentage of target bases covered at a minimum of 30X. It provides an estimate of the number of regions of interest that are sufficiently covered to ensure reliable analysis. Its lower limit is 99%.*

- Sequencing

-Q30_PCT: *This represents the percentage of bases sequenced with a quality score higher than 30, which reflects the accuracy of base calling. In this case, this accuracy is 99.99%. This parameter provides information on the overall run quality, as well as the quality of base calling and the optical signal used to read them. For example, Q30 of 95% means that 95% of the bases were sequenced with 99% accuracy. Its lower limit is 90%.*

-PF_BASES (Passing Filter bases): This is the total number of bases that passed the initial quality filters. This parameter verifies that the desired amount of raw sequencing data was achieved for each sample. In particular, it allows for the evaluation of library pooling (balancing) during multiplexing. Its lower limit is 2.30847×10^{11} bases.

-TOT_DUPLICATE: Reflects the percentage of duplicated reads (PCR or optical duplicates). When high, it can indicate a technical error or over-amplification during PCR. Since duplicated reads are redundant (non-independent artifactual observations), they are ignored during variant calling and therefore represent a loss of raw data. The upper limit is 27%.

- Read Alignment

-READS_ALIGNED_PCT: This is the percentage of reads aligned to the Hg19 reference genome through the Humanomics pipeline. This metric verifies the success of the alignment step within the Humanomics pipeline. The lower limit is 99.6%.

- Variant Calling

-SNP_REFERENCE_BIAS: *Measures whether reads tend to preferentially align to the reference allele rather than the alternative allele (alignment bias in favor of the reference allele). It is the average fraction of reads for which the reference allele is observed at a site where a heterozygous SNP is present. This parameter reflects the slight bias related to capture. An upward deviation is also indicative of contamination. The maximum limit is 0.55*.

3. Tertiary bioinformatics analysis: Franklin validation

Once the quality parameters were investigated for each sample, we proceeded with variant analysis and interpretation. To do so, the VCF files were imported into the Franklin software.

Initially, this analysis was performed using Alissa software. Following the discontinuation of this platform, we had to switch to the Franklin software. These are the differences:

- Alissa had a rather closed, restricted working environment with a well-defined workflow and precise databases, without direct interaction with the cloud. Franklin, on the other hand, is more open, being connected to the cloud. Consequently, it considers far more elements and databases than Alissa in its variant analysis.
- Alissa used fixed databases, annotation engines, and variant classification systems. Franklin, in addition to incorporating Alissa's tools, features an AI (artificial intelligence) function that allows it to gather the maximum amount of information from various databases and sources, with the aim of identifying variants as accurately as possible and assigning them the most appropriate classification using ACMG classification rules.
- The Alissa platform offered certain possibilities and limitations in the construction of sorting trees, which are not entirely identical in Franklin.

To investigate all these differences, we analyzed 25 positive samples and 25 negative samples on both platforms. The VCF files from the same samples were imported into both platforms to verify concordance in the detected variants, concordance in zygosity, concordance in variant classification, concordance in the number of detected variants, as well as concordance in the transcripts used by both platforms. The results will be presented in the Results section.

The list of genes targeted by BabyDetect for the targeted panel is as follows (see Appendix 1): the genes highlighted in yellow represent those associated with diseases currently screened for by the national newborn screening program in Belgium.

4. The computer software used

Throughout this work, we used computer software:

- GLIMS: allowed us to access the samples recorded in the system, to check that the data transfer systems from some equipment to the analysis Dashboard were working correctly. In particular, the transfer of data from the Puncher to the Dashboard, and the transfer of data from the Dashboard to GLIMS as well. The validation of this macro transfer has been formalized and published in Vivaldi document control QMS under the name BGE.DATA.INF.A03.

- Alissa-Agilent and Franklin-Qiagen: these variant analysis and interpretation platforms allowed us to gather sample data (genes, variants, zygosity, allele frequency, positions on the

transcript, protein) analyzed on this platform to compare them with data from the same samples analyzed on the Franklin platform.5. NGS-WES validation.

- VIVALDI: It's the electronic quality management system of the laboratory. We have two important modules. Vivaldi document control: It allowed us to update procedures, create new ones and publish validation plans. The Vivaldi process 2 Flow: allowed us to record any non-conformities that occurred during our work, update the risk analysis. It allows monitoring electronic traceability. Appendix 2 covers the actions and documents we've worked on and that have been published.

5. NGS-WES validation

5.1 Validation plan

The risk analysis for moving and post-moving also allowed us to develop a validation plan for the BabyDetect analysis by next-generation sequencing on filtered exome. This plan will be presented in the results section.

5.2 The NGS-WES Workflow – Material and Methods

5.2.1. Sample Reception and Preparation via Twist Technology

This step proceeded as described in the BabyDetect targeted NGS panel. The Kits used were:

- Enzymatic DNA Fragmentation: Performed using the Twist Library Preparation EF kit 1,2,0 v.2.0 (Twist Bioscience – ref: 104119).
- Adapter Ligation: With the Twist Universal Adapter System TrueSeq compatible Plate A (Twist Bioscience – ref: 101311)
- PCR Amplification: a PCR amplification stage was performed using primer pairs containing specific indices. The number of cycles here was 8 instead of 6 like with NGS-panel.
- Sample Pooling: multiplex pooling was performed combining 8 samples to yield a total of 1 pool. The objective was to achieve an approximately uniform mass concentration across all patients (250ng/patients).
- Hybridization and Capture: Twist Custom Panel (Twist Bioscience – ref: TE-98209850), Twist Universal Blockers (Twist Bioscience – ref: 100578) and Twist Exome 2.0 +Comp Exome Spike in (Twist Bioscience – ref: 105034).
- Washing and Enrichment: Post-capture PCR amplification was then carried out using a master mix Equinox Library Amp Mix (Twis Bioscience – ref: 104107).

5.2.2 Sequencing and Bioinformatics Analysis

This was performed using the NextSeq300 v2.5 HM kit (149-10-10-149) and the Illumina NextSeq 550 sequencer.

5.2.2.1 Bioinfomatic pipeline: secondary bioinformatic analysis

For bioinformatics analysis, we opted for a research use only commercial solution: Illumina's Dragen platform. It is still in development phase for the BabyDetect Test.

During a major DNA data analysis competition in 2020 (called the Precision FDA Truth Challenge V2), the DRAGEN[40] version 3.7 tool was recognized as the most reliable. It delivered good results, even in the most complex areas of the genome to analyze, and this applied to all types of Illumina sequencing data. The subsequent version, v4.4, also offers good results in terms of accuracy. DRAGEN pipelines can analyze several types of data: whole genome, whole exome, gene panels, single-cell RNA-Seq (i.e., to study gene expression one cell at a time), single-cell ATAC-Seq (to study DNA accessibility for each cell), bulk RNA-Seq (to study gene expression by groups of cells), and methylation.

Due to its reliability and speed, which have been demonstrated by researchers at Baylor College of Medicine[41], and as described above, we chose to use the Dragen platform here for the NGS filtered exome analysis. Dragen's architecture makes it possible to identify genetic variations, even in regions of the genome that are usually complicated to study. It can detect single nucleotide variants (SNVs); small base insertions or deletions (indels); copy number variations where a piece of DNA is duplicated or missing (CNVs) and large DNA rearrangements (structural variants). In the present case, we focused on CNVs, SNPs, and INDELs. To verify the pipeline's performance, we analyzed our data on this platform, the steps of which are as follows:

- Raw sequencing data retrieval: the sequencer is connected to Dragen (Dynamic Read Analysis for Genomics); sequenced data are directly generated and retrieved on the Dragen platform using Dragen BCL Convert.
- Read alignment: It is at this step, among others, that the difference is made compared to Humanomics. Dragen uses an algorithm with a field-programmable gate array architecture that each user can program: FPGA, which enables fast processing times. The initially compressed FASTQ files are decompressed and read; the pipeline evaluates the overall base quality, GC content, and read distribution, and aligns them to the Hg19/GRCh37 reference genome in our case. At the end of this step, BAM files are obtained, just as with Humanomics.

- Sorting, indexing, and variant calling: the aligned reads are sorted according to their chromosomal coordinates. They are then compared to the reference genome, and calculations are performed for each variant by considering certain parameters such as depth or allele frequency. The variants are then filtered based on these parameters. At the end of this step, we obtain VCF files, which are uploaded into the Franklin variant analysis software. The bioinformatics workflow is as follows (See Figure 4):

5.2.2.1Quality parameters: ee validation plan (Result section)

5.2.2.2 Validation parameters: see validation plan (Result section)

5.2.3 Tertiary bioinformatics analysis: see validation plan (Result section)

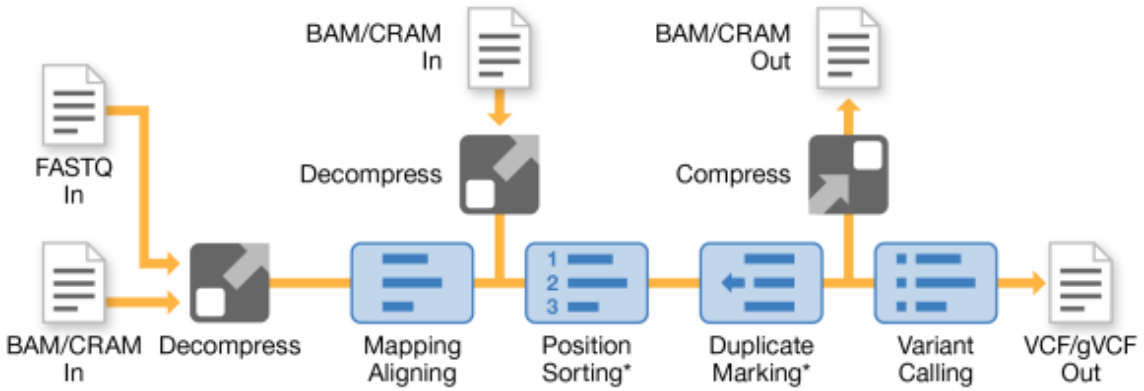


Figure 4: The bioinformatic pipeline: The compressed BAM files must be decompressed at each step where it is necessary for them to be read. Once the mapping, sorting, indexing, and variant calling are completed, the BAM file must be recompressed and the VCF file produced (from Illumina MD. *DRAGEN[39] secondary analysis)*

III. RESULTS

1.Risk analysis

This risk analysis considers the moving phase and the post-moving phase (NGS-panel and NGS-WES).

It highlighted the following points (Table 1):

			Moving risk analysis					
Identified risk	P	D	C	S	IC	Risk management	F	RC
Physical damage (to equipment/small material: (fall, impact))	Low	High	Shutdown or malfunction of equipment/small tools	Medium	Medium	Establishment of a relocation plan Backups are also planned in case of malfunction or absence of the current automaton.	Easy	Low
Loss of stock tracking sheets present on the refrigerators and freezers	Low	High	Loss of data	Medium	Medium	Gather the files one day before the move into a binder and transport them with the other binders. Create a digital version that reproduces the current manual data of the files	Easy	Low
Poor management of the current stock: reagents/samples/equipment moved for the new site, consumables (loss or otherwise)	Low	High	Waste of time, waste of money (in the case where we recommend what we already have)	Medium	Medium	Make an inventory of equipment, reagents, and samples in the refrigerators and freezers and in stock before the move, and make a comparison after the move in terms of new location in order to know exactly where to find each item. Conduct a visit of the new premises to identify and confirm the new destination of each piece of equipment, consumable, reagent, or sample. The old stocks of reagents or samples kept in the cold will be used up or discarded after the move if they are no longer needed.	Easy	Low

Table 1: Risk analysis (moving and post moving): The IC column shows the level of criticality of the identified risks considering the measures already in place. The RC column indicates the level of risk that is desired once all the measures proposed in the 'Risk management' column are implemented and the risk analysis is reassessed. Column C = Consequences.

	Postmoving risk analysis								
	Identified risk	P	D	C	S	IC	Risk management	F	RC
Workspace	Wokspace management: are they available at the time of the analysis ? Yes ,no Equipment availability	Medium	High	longer TAT	Medium	Medium	Reservation of Bench/equipment/small tools and ensuring their availability for the analysis we want to perform	Easy	Low
Staff	Less controlled air flow:entering and exiting more frequent than before: risk of contamination	Low	Medium	Unusable result	Medium	Medium	Training on laboratory best practice rules: Limiting entries and exits during manipulations, avoiding contact with people in the post area when still in the pre area (contact in the corridors, for example). Respect the pre-post flow by making sure to change clothes, keep aprons in the rooms before leaving, and put them back on when entering the rooms.	Easy	Low
ANA	Risk of contamination by surfaces/ shared pipettes or any other equipment or consumables	Low	High	Unusable result	Medium	Low	Training on: Cleaning benches, pipettes, and other equipment after each handling for the person leaving, and if possible before each handling for the person using the same workstation afterwards.	Easy	Low
	Risk of sample reversal	Low	Medium	incorrect result	High	Medium	Accepted risk Un second tier test is planned in case of positive result: Sanger sequencing or any other method	NA	NA
	Risk of mixing between Samples/reagents Risk of contamination in shared refrigerators	Low	High	No result or unusable result	Medium	Low	Aliquoted and well-labeled reagents before placing them in shared refrigerators In each shared refrigerator, a label is placed on the drawer indicating the people who have access. E.g.: "Microbio-Babydetect." A physical barrier between samples and reagents is created inside the drawer of the shared fridge or freezer: Reagents are placed in a box labeled with the nature of the contents and the owner. E.g., "REAGENTS - BabyDetect" on one box and "PRODUCTS - Babydetect" on another box, both stored on the left side of the drawer, leaving, for example, space on the right for the other sector using the drawer and handling its samples and products according to the same protocol.	Easy	Low
	Samples drop: Los of samples or any other incident during transport from the laboratory to sequencer room	Low	High	No result, replanification of the analysis	Medium	Low	Sample transport devices on ice planned, sufficiently spacious; Eppendorfs/tubes containing the samples well closed before movement	Easy	Low
	Difficulties to read the reference number of the flow cell on the sequencer	Low	High	Impossibility to run the analysis	Medium	Low	Flow cells planned in stock and make sure that the stock is okay before starting the analysis.	Easy	Low
Preana	Delay in Goldcard encoding/analysis process/transmission of results	Low	High	longer TAT	Medium	Low	Backup planned, to be trained. Training evidence to be added to the operator's skills file.	Easy	Low
Franklin	Discordance in the inferred variants when changing platform. Differences in variant coverage during screening by the NGS-exome method	Low	Medium	Wrong result Communication d'un résultats inexistant	High	Medium	Validation plan	Medium	Low
Stock	Out of Stock	Low	High	Longer TAT	Medium	Low	Preparation of a package of necessary consumables before starting a procedure and ensuring that the stock is still sufficient for another procedure	Easy	Low

2. Revalidation following relocation: validation plan and results

Right from the outset, we recall here that all BabyDetect's current activities focus on next-generation sequencing (Targeted panels and whole exome), within the framework of screening for rare neonatal pathologies whose manifestations occur early in childhood: within the first 5 years. These activities encompass all the steps necessary for their execution, including DNA extraction, library preparation, as well as the sequencing itself.

As part of the platform's relocation, we proceeded in two stages. The first phase focused on the relocation of the extraction stage. This move was scheduled for the week of March 16, 2026. The second phase focused on the relocation of the remaining workflow.

To prepare for the relocation, a meeting was held on February 17, 2026, from 10:30 am to 12:00 pm, both with the managers of the current premises and with the managers of the various sectors of activity using the premises of Tower 6, +1 (new premises), all under the supervision of the quality manager. In the same vein, a meeting was also held on February 26 with the storage manager of Unilab-lg. Following these various meetings, the objective was to formalize everything that had been discussed through a risk analysis and this validation plan.

2.1 Phase 1: Revalidation of the extraction step

Objective: To demonstrate that Nucleocube1 functions identically from a mechanical, electrical, and IT perspective after the relocation.

Implementation: To limit the risks of impact and dropping of the automated system, the CHU handling team was contacted, via the logistics foreman (Mr. Marc Baade), to transfer this equipment from the current room to its future position in the new premises. On February 25, 2026, two handling agents visited the current site to inspect the equipment and technically plan its transport. The relocation of the Nucleocube1 was carried out on March 17, 2026, along with the relocation of the small equipment: plate sealer, plate vortex, and thermomixer.

In tandem, Mr. Jean-Luc Deladrière, manager at the Biocubik firm, was also contacted to set a date for the requalification of the Nucleocube1; this was scheduled for March 20, 2026, and the requalification effectively took place on that date.

Results:

2.1.1 Nucleocube 1 requalification: The report from Biocubik's engineer, the supplier of the Nucleocube, shows a positive result for the installation requalification and operational requalification.

The re-qualification report indicates an absence of damage to the equipment and its proper functioning with respect to the expected specifications. The validation reports entitled: *qualification report Nucléocube 1_1.1* and *qualification report Nucléocube 1_1.2* can be

consulted in Vivaldi P2F / BGE.BD/NC/01 under the document tab. Each logbook associated with equipment has a unique code.

2.1.2 DNA concentration comparison before and after relocation: The average DNA concentrations measured after extraction before and after the move differ significantly. A significant difference is also observed when comparing the average concentrations obtained between two runs before the move.

In addition to operational requalification, we performed a comparison of DNA concentrations measured using the Qubit BGE.BD/FL/01 before and after the move, between run 94 before the move and run 115 after the move (first run whose extraction was done after the move, the library preparation having been done in the GIGA premises). The comparison was performed using Welch's t-test and the Mann-Whitney test for independent samples using R software version 4.3.1.

The mean DNA concentration obtained before the move was 4.01 ng/µl and the median was 3.85 ng/µl. The mean DNA concentration obtained after the move was 2.99 ng/µl and the median was 2.90 ng/µl. These two values are not far apart from each other in the two groups.

The histograms of DNA concentrations before and after, the quantile-quantile plots in both groups, the Shapiro-Wilk test and the homoscedasticity test allowed us to conclude that the data in these two groups follow a normal distribution. However, the difference in variance between the two groups led us to apply Welch’s test.

Welch's t-test for independent samples yielded a p-value <0.0001, indicating a highly significant difference between the DNA concentrations obtained before (mean = 4.01 ng/µl) and after (mean = 2.99 ng/µl) the move.

To further investigate this result, we also performed a Mann-Whitney test, which showed a significant difference ($p<0.0001$) between the two medians (before = 3.85 ng/µl and after = 2.90 ng/µl) before and after the move.

To investigate whether the difference was due to a run effect or to the move itself, we compared run 96 before the move with the same run 115 after the move. Using the same methodology as above, we assumed normal distribution in both groups with a difference between variances. Welch's test gave a p-value of 0.9, supporting equality of the mean DNA concentrations before (mean = 3 ng/µl) and after (mean = 2.99 ng/µl) the move.

We also compared pre-move runs with each other (Figure 5): run 96 before the move versus run 94 before the move. Using the same methodology as above, we assumed normal distribution in both groups with a difference in variance between the two pre-move runs.

Welch's test gave a P value <0.0001, supporting a difference in DNA means: mean run 96 = 3 ng/µl and mean run 94 = 4.01 ng/µl.

The Mann-Whitney test for independent samples confirmed this result with a p-value <0.0001, supporting a significant difference in DNA means between the two pre-move runs (median run 96 = 3.06 ng/µl vs median run 94 = 3.85 ng/µl).

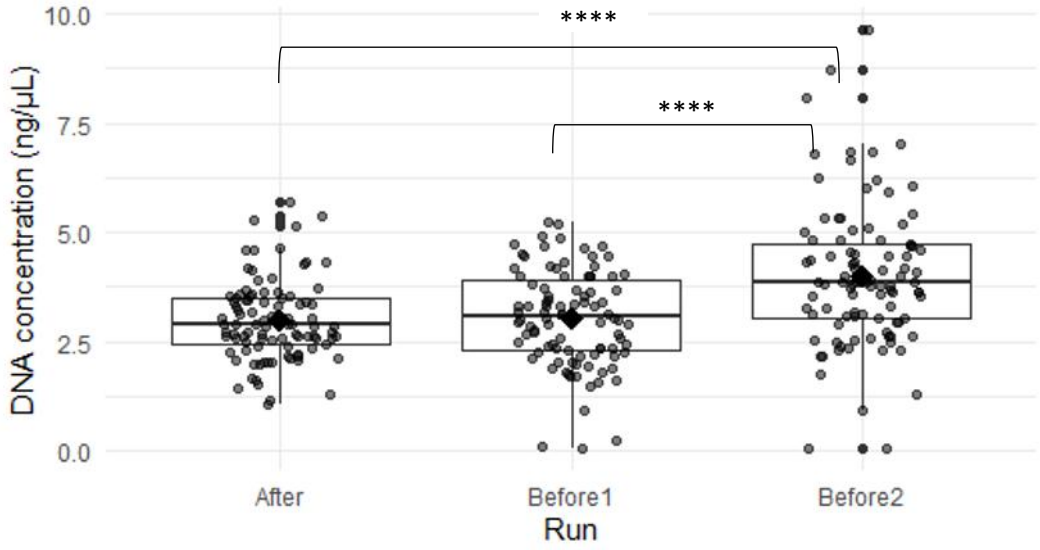


*Figure 5: Extracted DNA concentration Before and after relocation: distribution of DNA concentrations measured after extraction, for each sample in runs 96 and 94 before relocation, as well as in run 105 after relocation. On the x-axis. before 1 = Run 96, Before 2 = Run 94 and After = Run 115, and on the y-axis, we have the DNA concentration value in ng/µl. (*** = statistically significant difference).*

The observed differences are due to inter-run variability and not directly to the move or analytical variability. The minimum concentration expected for the analysis to be proceed is 1ng/µl. All the samples before and after have concentration above the threshold.

2.1.3 DNA fragment size comparison: the fragment sizes are within the expected values

We also carried out a comparison of the size of DNA fragments after extraction, after the move. We used the fragment analyzer, and the comparisons were made between the fragment sizes for each of the 23 samples and the size we expected to have.

We expected a size of at least 30 kb after DNA extraction. Of the 23 samples analyzed after the move, 18 had a size greater than 60 kb and 5 had an average size of 37 kb.

2.2 Phase 2: Revalidation of the Remaining Targeted NGS Panel Workflow

The relocation of the remaining workflow was progressively carried out during the month of May 2026, including the relocation of the thermocyclers, as well as the reagents and

consumables required for library preparation through sequencing.

2.2.1 Thermocycler Verification

As part of the relocation process, the new laboratory environment offered the possibility to use different thermocyclers, unlike the previous setup where the same thermocyclers were consistently used. Therefore, both the relocated thermocyclers and the thermocyclers already present in the new facility were used.

To demonstrate equivalence between the thermocyclers, an additional section entitled Thermocycler Robustness was added to the validation file under the robustness parameter.

Objective: To verify that the thermocyclers operate equivalently and provide the expected results in the new laboratory environment.

Implementation: The internal verification was performed using Driftcon, software version 4.5 from Cycler test supplier, by measuring the thermocycler temperature. The check consisted of verifying that the thermocycler produces the expected temperatures for the chosen cycles. The measured values were compared with the reference values.

Results: The thermocycler BGE.BD/TH/01 was verified in the pre-PCR area on 29/04/2026 after installation in the new laboratory. The validation report indicates that the thermocycler remains capable of providing the expected results. This report is available in Vivaldi P2F within the equipment logbook. The thermocycler BGE.BD/TH/02 was also verified.

2.2.2 Revalidation of library preparation for sequencing

Objective: Check that we are still able to adequately detect variants after preparing the samples with all the equipment and facilities in the new location.

Implementation: We carried out several tests which we will describe and for which we provide the results below.

Since the sequencers have not changed location, a good sequencing result will indicate proper preparation of the libraries in the new facilities. Overall, the results of samples sequenced after the move are compared to those obtained before the move.

Results:

2.2.2.1 Pre and post capture DNA comparison before and after relocation: The average concentrations of pre-capture DNA libraries before and after the move show a statistically significant difference. The average concentrations of post-capture DNA pools before and after the move do not show a statistically significant difference.

We compared graphically and statistically the DNA quantification pre- and post-capture before the move for RUN 105 and after the move for RUN 116 (first run where the full BabyDedect analysis workflow was completed on Unilab's premises). For this, we performed a Mann-Whitney test for independent samples and a student's t-test for independent samples using R software version 4.3.1.

- **Statistical Comparison of Library Means (Pre-capture)**

The mean library concentration obtained before the relocation is 214.90 ng/µL, with a median of 220 ng/µL. The mean library concentration obtained after the relocation is 124.71 ng/µL, with a median of 119.5 ng/µL. In both groups, these two values (mean and median) are close to each other.

The histograms of DNA concentrations before and after the relocation, the quantile-quantile (Q-Q) plots for both groups, the Shapiro-Wilk test and the homoscedasticity test allow us to conclude that the data in both groups follow a normal distribution with equal variances.

A Student's t-test yielded a p-value < 0.0001, indicating a highly significant difference between the library concentrations obtained before (Mean = 214.9 ng/µL) and after (Mean = 124.71 ng/µL) the relocation.

- **Statistical Comparison of Post-capture Pool Means**

The mean pool concentration obtained before the relocation is 2.49 ng/µL, with a median of 2.57 ng/µL. The mean pool concentration obtained after the relocation is 2.84 ng/µL, with a median of 2.60 ng/µL. These two values are far apart in one of the two groups.

The histograms of DNA concentrations before and after relocation, the quantile-quantile (Q-Q) plots for both groups and the Shapiro-Wilk allow us to conclude that the data do not follow a normal distribution in either group. Consequently, we performed a Mann-Whitney test, which yielded a p-value = 0.42, indicating no significant difference between the pool concentrations obtained before (Median = 2.57 ng/µL) and after (Median = 2.60 ng/µL) the relocation.

The differences observed in pre-capture DNA concentrations are potentially due to differences in DNA concentrations after extraction, which can also be due to the initial sample. During pool preparation, the library concentrations to be factored in per pool are equivalent (187 ng/µL), which justifies the equality of post-capture pool means.

2.2.2.2 Longitudinal follows up of quality parameters for Internal quality control sample before and after relocation: The internal control after the move meets all the quality criteria for all the quality parameters represented

We conducted a longitudinal follow-up of the quality parameters of internal controls across all runs performed by the NextSeq sequencers before and after the relocation. The metrics studied were as follows: %Q30, Mean Coverage, % Read Aligned, % Target Bases 30X, % Duplicates, and SNP Reference Bias.

The acceptance thresholds (See figure 6) set for these different parameters were respectively: %Q30: ≥ 90%, Mean Coverage: ≥ 117X, % Read Aligned: ≥ 99.6%, % Target Bases 30X: ≥ 99%, % Duplicates: ≤ 27%, and SNP Reference Bias: ≤ 0.55.

Analysis of the longitudinal follow-up curves shows that all quality parameters of the internal controls remain within the acceptability limits defined after the move. These results support the maintenance of the analytical performance of the sequencing process following the transfer of the activity.

2.2.2.3 Comparison of analytical variation before and after relocation: The analytical variability is the same before and after relocation

The following quality metrics were compared (see figure 7) between the samples from run 114b carried out before the move and those from run 116 carried out after the move: %Q30, Mean Coverage, % Read Aligned, % Target Bases 30X, % Duplicates, and SNP Reference Bias. This analysis was conducted using Levey-Jenning's type longitudinal tracking charts generated in RStudio version 2026.01.2+418 as well as statistical tests carried out with R version 4.3.1.

Overall, differences in values between samples were observed for several quality parameters between the two runs. However, all values remain within the acceptability thresholds defined for analytical validation. To statistically evaluate these differences observed graphically, comparisons were made between the samples of runs 114b and 116 for each of the parameters studied.

Graphical analysis of the histograms showed a slightly skewed distribution before relocation and a more symmetric one after relocation. The quantile-quantile plots (Q-Q plots), the Shapiro-Wilk test and the variance homogeneity test allow us to conclude to a normal distribution. Given these elements, a comparison of means was performed using an independent samples Student's T-test. A statistically significant difference in the average coverage of the target regions was observed between runs 114b (average = 285.38X) and run 116 (363.87X): $p < 0.0001$

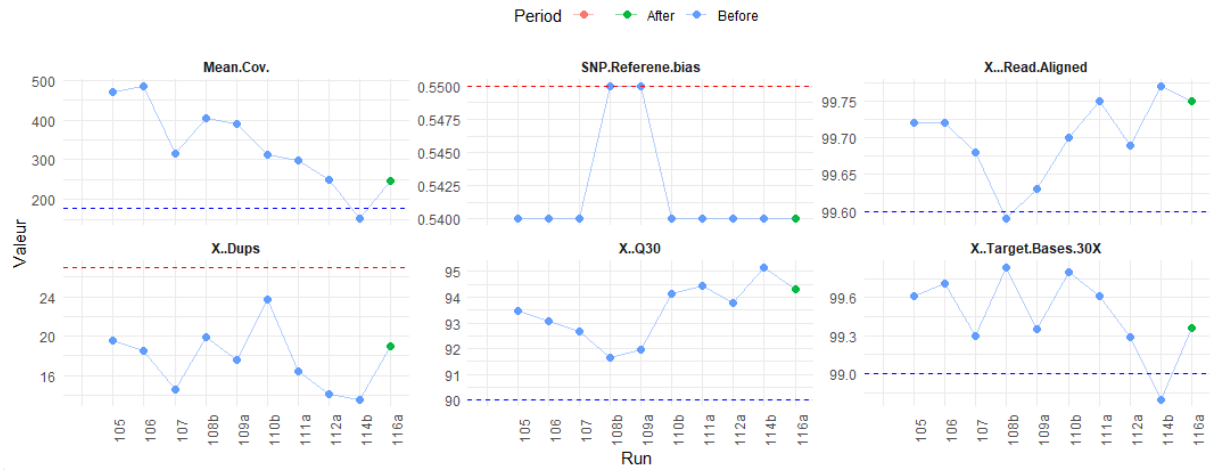


Figure 6: Longitudinal follows up curve of quality parameters for Internal quality control sample before and after relocation: The horizontal lines represent the minimum quality thresholds required and the maximum threshold not to be reached for a sample to be considered of good quality or to meet the quality criteria. On x-Axis we have different runs and on the y-axis the value of the measured parameters. All the quality criteria are met after relocation.

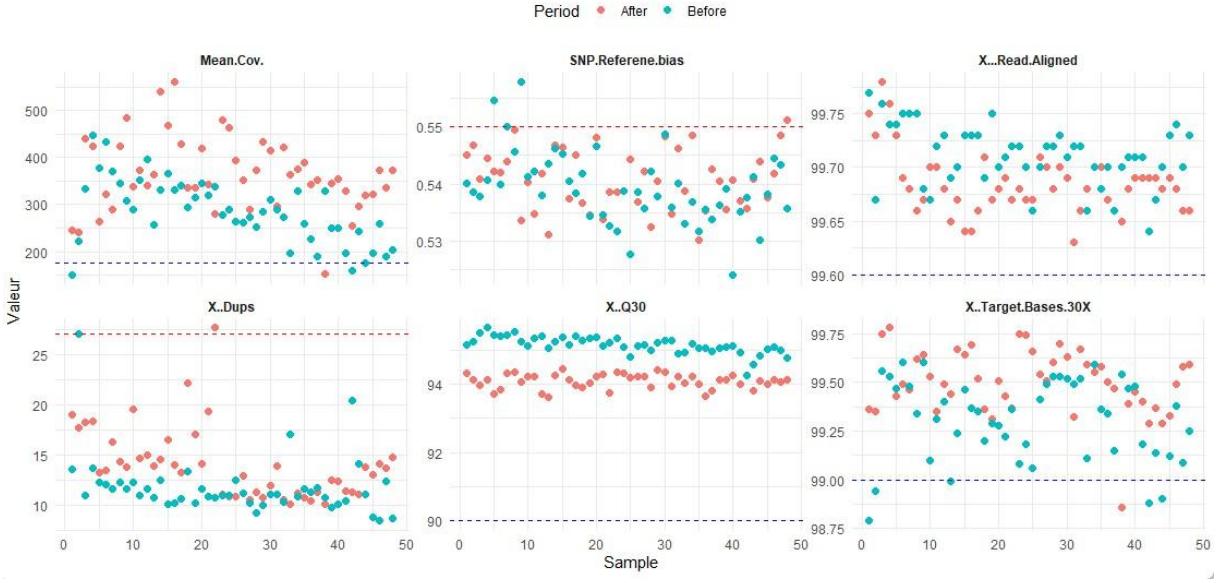


Figure 7: Longitudinal follow-up curves of quality parameters for samples from runs 114b (before the move) and 116 (after the move): The horizontal lines represent the minimum quality thresholds required and the maximum threshold not to be reached for a sample to be considered of good quality or to meet the quality criteria. From left to right, each point represents independent samples from 1 to 48 in both runs. A comparison is made between the observed values for independent samples before and after the move, considering defined parameters. The overall distribution of the samples before and after the move is identical.

Nevertheless, all observed values remained well above the minimum acceptability threshold set at 117X.

- **For the “SNP reference bias" parameter:**

After relocation (run 116): The mean for SNP reference bias was 0.54 with a median of 0.54. Before relocation (run 114b): This value was 0.53 with a median of 0.53. The means and medians observed in both groups were identic, suggesting an overall symmetric distribution.

Graphical analysis of the histograms showed a symmetric distribution. The quantile-quantile plots (Q-Q plots), the Shapiro-Wilk test and the variance homogeneity test support a normal distribution in both groups.

A Student's t-test for independent samples was therefore applied. No statistically significant difference in SNP Reference Bias was observed between the runs before and after the move ($p = 0.2$).

These results support the stability of the reference bias after moving. The remained well below the minimum acceptability threshold set at 0.55.

- **For the “% Read aligned" parameter:**

After relocation (run 116): The mean for % read aligned was 99.68% with a median of 99.68%. Before relocation (run 114b): This value was 99.71% with a median of 99.71. The means and medians observed in both groups were identic, suggesting an overall symmetric distribution.

Graphical analysis of the histograms showed a symmetric distribution. The quantile-quantile plots (Q-Q plots), the Shapiro-Wilk test yielded, and the variance homogeneity test support a symmetric distribution.

Given the large sample size and the proximity between means and medians, a student’s T-test for independent samples was selected.

A statistically significant difference in the percentage of aligned reads was observed between the runs before (mean = 99.71%) and after (mean = 99.68%) the move ($p < 0.0001$).

However, the observed values remained above the acceptability threshold set at 99.6%.

- **For the “% Duplicate" parameter:**

After relocation (run 116): The mean for % Duplicate was 13.98% with a median of 13.57%. Before relocation (run 114b): This value was 11.76% with a median of 11.05. The means and medians observed in both groups were not close, suggesting an overall asymmetric distribution.

Graphical analysis of the histograms showed a symmetric distribution. The quantile-quantile plots (Q-Q plots) and the Shapiro-Wilk test yielded us to a non-symmetric distribution.

Considering these elements, a non-parametric comparison was performed using a Mann-Whitney test for independent samples.

We obtained a significant difference between the %duplicate in the Runs before (median = 11.05%) and after (median = 13.57%) with a p value <0.0001.

However, the observed values remained below the maximum acceptability threshold set at 27%.

- **For the “% Q30" parameter:**

After relocation (run 116): The mean percentage of bases sequenced with a quality ≥ Q30 was 94.09% with a median of 94.12%. Before the move (run 114b), the mean was 95.14% with a median of 95.15%. These two values are very close to each other in both groups.

Graphical analysis of the histograms showed an asymmetric distribution. The quantile-quantile plots (Q-Q plots), the Shapiro-Wilk test and the variance homogeneity test yielded to a normal distribution.

A parametric comparison using the Student's T-test for independent samples was performed.

We obtained a significant difference between the %Q30 in the Runs before (mean = 95.14%) and after (median = 94.09%) with a p value <0.0001. However, the observed values remained above the minimum acceptability threshold set at 90%.

- **For the “Target bases 30X" parameter:**

After relocation (run 116): The mean percentage for Target bases 30X was 99.50% with a median of 99.50%. Before relocation (run 114b): This value was 99.30% with a median of 99.34. The means and medians observed in both groups were close, suggesting an overall symmetric distribution.

Graphical analysis of the histograms showed an asymmetric distribution. The quantile-quantile plots (Q-Q plots) showed points aligned along the line in one group compared to the other group and the Shapiro-Wilk test yielded values in favor of asymmetrical distribution in both groups.

Considering these elements, a non-parametric comparison was performed using a Mann-Whitney test for independent samples.

A statistically significant difference was observed between the runs before (median = 99.34%) and after (median = 99.50%) the relocation: $p < 0.0001$.

Nevertheless, all observed values remained well above the minimum acceptability threshold set at 99%.

To assess the analytical variability before and after the move, we calculated the coefficients of variation for each parameter before and after the move. The results are recorded in the table below (Table 2):The coefficients of variation after and before moving are of the same order of magnitude, which suggests that the variability before and after is controlled.

2.2.2.4 Comparison of metrics before and after the move for identical samples: The samples that already met the quality criteria before the move also meet them after the move

The quality metrics below (see Figure 8) were graphically compared between identical samples before (14-00BD00-0101: run 101, 14-210000-0105: run 94, 14-210000-0106: run 94 and 14-25BD21-005201 run 101) and after (14-00BD00-0116: run 116, 14-210000-010501:run 116, 14-210000-010502:run 116, and 14-25BD21-005202:run 116) relocation. QC metrics: %Q30, Mean Cov, % Read Aligned, % Target Bases 30X, % Dups, and SNP reference bias. This was performed using RStudio software version 2026.01.2+418."

Overall, variations in quality metrics were observed between the pre- and post-relocation analyses. However, with one exception, all values remained within the defined acceptability thresholds. These results suggest an analytical variability compatible with the inter-run variability usually observed in routine practice, with no major impact on the overall performance of the analytical process.

2.2.2.5 Comparison of the detected variants before and after relocation: The true variants detected by the variant analysis and interpretation software before the move are the same ones detected after the move. The concordance of variants in the VCFs of the same samples before and after the move is over 90%.

To verify that the relocation had no impact on variant detection and the generation of bioinformatics output files, the VCF files obtained before and after the relocation were compared for the same set of samples. This comparison was performed using the RTG Tools software to evaluate the concordance of detected variants with a sequencing depth of 30X and located within the BabyDetect regions of interest.

The analysis also focused on variants located within the BabyDetect regions of interest

Parameters	Period	N	Mean	SD	CV
X..Dups	Before	48	11.7639583	3.005101152	25.54498296
X..Dups	After	48	13.9850000	3.528893504	25.23341798
Mean.Cov.	Before	48	285.3816667	69.074216764	24.20415354
Mean.Cov.	After	48	363.8704375	77.914876737	21.41280761
SNP.Referene.b	Before	48	0.5392610	0.006324085	1.17273179
SNP.Referene.b	After	48	0.5407613	0.005269265	0.97441608
X..Q30	Before	48	95.1458333	0.253845600	0.26679634
X..Q30	After	48	94.0902083	0.208535824	0.22163393
X..Target.Bases	Before	48	99.3056250	0.210913981	0.21238876
X..Target.Bases	After	48	99.5008333	0.165642955	0.16647394
X...Read.Aligne	After	48	99.6847917	0.029462532	0.02955569
X...Read.Aligne	Before	48	99.7108333	0.028941124	0.02902502

Table 2: Comparison of the analytical variability before and after the move: SD represents the variance associated with the mean, and CV represents the coefficient of variation. For each parameter, the variability before the move is compared to the variability after the move between two runs.

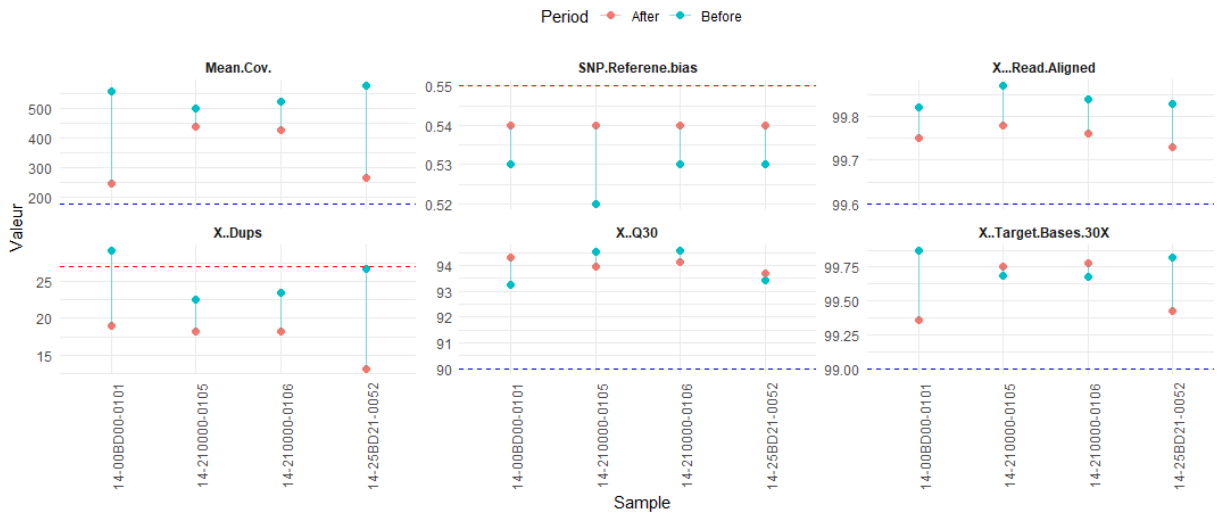


Figure 8: Longitudinal follow-up curves of quality parameters for identical samples before and after the move: *The horizontal lines represent the minimum quality thresholds required and the maximum threshold not to be reached for a sample to be considered of good quality or to meet the quality criteria. On x-Axis we have the 4 samples and on the y-axis the value of the measured parameters.*

intersecting the high-confidence regions defined in the GIAB HG002 reference material BED file.The concordance between the VCF files was calculated (see Table 3) using the following formula: C / (A + B + C). Where:

-C corresponds to the number of variants identified in both compared VCF files;

-A corresponds to the number of variants identified in the first VCF file but absent from the second;

-corresponds to the number of variants identified in the second VCF file but absent from the first.

GLIMS	Compa_ROI_DP30	Compa_ROI_HCR_DP30
14-00BD00-010101-vs-14-00BD00-011601	0,904441	0,933333
14-00BD00-010101-vs-14-00BD00-011602	0,902291	0,931085
14-00BD00-011201-vs-14-00BD00-011601	0,903872	0,925764
14-00BD00-011201-vs-14-00BD00-011602	0,906836	0,930556
14-210000-0105-vs-14-210000-010501	0,881242	0,911011
14-210000-0106-vs-14-210000-010502	0,877883	0,901124
14-25BD21-005201-vs-14-25BD21-005201	0,905559	0,926829

Table 3: VCF concordance between identical samples before and after relocation: Compa_ROI_DP30 = comparison of variants with a sequencing depth of 30X within the BabyDetect regions of interest, Compa_ROI_HCR_DP30 = comparison of variants with a sequencing depth of 30X in the BabyDetect regions of interest intersecting the high-confidence regions defined in the GIAB HG002. HCR are all the regions of the GIAB genome that have been fully characterized. They cover about 85% of the genome (Hg19) and define the regions that can be used as a reference standard.

between 80 and 90% is considered good. The discordances are mostly due to false positives, which will be highlighted by the Tri Franklin tree ("Variant Allele Frequency (VAF)" node that removes SNPs with an allele frequency between 0.0 and 0.3).

For an internal control (14-00BD00-0101 before vs 14-00BD00-0116 after) as well as a positive sample (14-25BD21-005201 before vs 14-25BD21-005202 after), the coverage, gene, zygosity and variant allele frequency were compared before and after the relocation.

Sample	Gene	Variant	Zygosity	VAF	Period
14-25BD21-(005201vs 02)	FAH	c.554-1G>T	homozygote	0.02	Before
	FAH	c.554-1G>T	homozygote	0.02	After
14-00BD00-(0101vs 0116)	/	/	/	/	Before
	/	/	/	/	After

Table 4 : Comparison of variants identified in the same samples before and after relocation

These results support the maintenance of variant detection and interpretation performances after the relocation.

2.3 Robustness

The sample preparation for run 115 was done using the thermocyclers BGE.BD/TH/01 and BGE.BD/TH/02 at the GIGA premises, and that for run 116 was done with the thermocyclers at Unilab's premises: BMO/TH/01, BMO/TH/08, and BMO/TH/30. Those thermocyclers used during library preparation in Unilab's premises were checked using a Driftcon (see 2.2.1). The verification reports support the proper functioning of each of these thermocyclers by producing the expected temperatures. These reports can be consulted in the logbook of each of these pieces of equipment on Vivaldi P2F.

The satisfactory results obtained using this equipment, which are in line with the expected results, demonstrate the equivalence between the different thermocyclers and thereby validate the use of one thermocycler or another within the Pre-PCR, PCR, and Post-PCR room areas.

3. Validation of the Franklin Interpretation Software

Across the 50 samples for which we compared the number of variants detected per identified gene using Alissa versus Franklin, we obtained the following results for each of the 25 positive samples and each of the 25 negative samples.

	Alissa	Franklin	Alissa + Franklin
Gens	25	25	25
Variants	37	37	37

Table 5: Concordance between variants detected with Franklin and Alissa: The intersection of gene + Alissa represents the number of genes found in Alissa:25 genes for the 25 positives samples.The intersection of variants + Alissa represents the number of variants found in Alissa (37). The same interpretation is applied to the intersection of the rows with the Franklin and Franklin + Alissa columns.

Franklin platforms. Among the genes identified by both platforms, 100% of the variants were detected by both software programs.

4. NGS-WES Validation: Validation Plan

The validation process was initiated with a design phase focused on the various regions of interest to be targeted. Discussions were conducted with the Twist bioinformatics team during several meetings, notably on January 27, April 15, and April 22, 2026.

A BED file was generated from version V2 of the PANEL, additionally incorporating synthetic regions of interest (spike-ins). This file was transmitted to Twist to produce probes required to capture the target regions of interest.

Several collaborative meetings took place regarding this matter between the BabyDetect team and the Twist team. Once manufactured, the reagents were delivered, followed by the testing phases.

Concurrently, discussions were held with Illumina regarding the Dragen secondary bioinformatics analysis. This platform was selected for this validation as the bioinformatics analysis tool due to its accuracy, speed, functionalities.This implies a reduction in processing time for the BabyDetect analysis. Discussions were initiated and pursued across various meetings, specifically on April 16, April 23, and May 12, 2026.

4.1The Testing Phase

This phase was carried out on 8 samples: 4 positive samples, 2 negative samples, 1 control sample, and 1 GIAB (Genome in a Bottle) in order to evaluate the initial performance of BabyDetect-spike-in-Exome, before moving to full validation

The workflow and the kits are identical to those of the targeted NGS panel workflow (BGE.BABYDETECT.ANA.06/section 2.2-Material and method), with a few exceptions described in section 4.2.1-Material and method.

For this phase, 50 ng of DNA were utilized for the 8 samples. During this phase, Twist was contacted, notably on May 27, to verify that the values obtained across the different steps of the workflow were satisfactory.

To integrate the Hg19-to-Hg38 transition into the BabyDetect workflow, a new BED file will be generated by Twist to verify that the biochemistry remains consistent with the targeted regions when considering Hg38 transcripts.

4.2 The Secondary Testing Phase

Four positive samples will be prepared using 40 µl of DNA to evaluate library preparation, applying 6 PCR cycles post-adapter ligation instead of the 8 cycles previously used. This will enable the assessment of the impact of cycle numbers and the determination of the optimal PCR cycle parameters to retain.

4.3 Validation

Validation will be performed on a half-plate of 48 samples using the NovaSeq 6000 S1 Reagent Kit v1.5 (200 cycles), considering the samples detailed below (Table 5):

grid - plate 1	1	2	3	4
A	**MV1.1-positive SNP**	**MV1.2-positive SNP**	**reference- HG002.1_NA24385- 50 ng**	**reference -HG002.2_NA24385 - 50 ng**
B	**MV2.1-positive SNP**	**MV2.2-positive SNP**	**MV1.1-negative**	**MV1.2-negative**
C	**MV3.1-positive SNP**	**MV3.2-positive SNP**	**MV2.1-negative**	**MV2.2-negative**
D	**MV4.1-positive SNP**	**MV4.2-positive SNP**	**MV3.1-negative**	**MV3.2-negative**
E	**MV5.1-positive CNV**	**MV5.2-positive CNV**	**MV4.1-negative**	**MV4.2-negative**
F	**MV6.1-positive CNV**	**MV6.2-positive CNV**	**MV5.1-negative**	**MV5.2-negative**
G	**MV7.1-positive CNV**	**MV7.2-positive CNV**	**MV6.1-negative**	**MV6.2-negative**
H	**MV8.1-positive CNV**	**MV8.2-positive CNV**	**MV7.1-negative**	**MV7.2-negative**

Table 6: NGS-WES validation plate: MV 1-4-positive SNP = 4 positive samples for SNPs in duplicate, MV 5-8-positive CNV = 4 positive samples for CNVs in duplicate, reference-HG002-NA24385-50ng = GIAB in duplicate with 50 ng DNA and MV 1-7-negative = 7 negative samples in duplicate.

The protocol will be followed as previously described, incorporating the PCR cycles determined during the testing phases. A volume of 40 µl of DNA will be utilized.

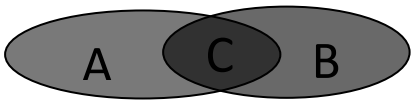
4.3.1 Validation Parameters

The GIAB will be used to determine:

- Sensitivity: which evaluates the probability of a true positive outcome among affected subjects. It will be calculated as follows: (Number of true positive / number of true positive + number of false negatives))
- Precision: which is the probability of truly having the condition when the test is positive. It will be calculated as follows: (number of true positives/ (number of true positives + number of false positives))

The positive, negative, and GIAB samples will be used to measure the concordance between the variants detected intra- and inter-run by the Dragen and Franklin analyses.

This concordance will be evaluated using the RTG Tools software according to the following formula: $\frac{C}{A+B+C}$



Where:

- C corresponds to the variants inferred in both compared replicates.
- A corresponds to the variants inferred in the 1st compared replicate but not in the 2nd replicate.
- B corresponds to the variants inferred in the 2nd compared replicate but not in the 1st replicate.

Repeatability and intermediate precision could also be evaluated.

4.3.2 Quality Parameters

The following quality parameters:Targeted_base_30X_pct ; Q30_pct ; SNP_reference_bias ; PF_bases ; Mean_target_coverage ; Read_aligned_pct ; Selected_bases_pct ; Mean_insert_size ; Tot_duplicates_pct ; Fold_80_base_penalty, currently monitored within the scope of the NGS-PANEL will also be monitored within the scope of the NGS-WES. Thresholds will be defined considering the distribution of our results obtained with the Dragen software.

Tertiary analysis will be performed using the Franklin interpretation software.

IV. DISCUSSION

1. Risk analysis and (re)validation plans

To carry out this work, a qualitative risk analysis was conducted, identifying the critical points requiring special attention to prepare for the revalidation of the BabyDetect analysis following the relocation, as well as after the change in interpretation software and analytical method. This analysis highlighted a risk related to facility management, namely the unavailability of a workstation at the desired time for processing, which would result in a longer Turnaround Time (TAT). Corrective actions were proposed to prevent this risk. The identified risk, along with the associated measures, were submitted to the administration responsible for deciding on their implementation. At this stage, the risk was reassessed and initially underestimated. Since the risk subsequently materialized, it was reconsidered, and an action was later implemented, specifically, the introduction of a booking schedule designed to avoid scheduling any processing when no space is available. Despite this measure, the risk occurred again.

This illustration highlights the importance not only of the risk analysis itself, but even more so of the involvement of decision-makers and, above all, operators (in this case, laboratory technicians) in the development, implementation, and compliance with preventive measures. It also sheds light on the need for rigorous monitoring of the actions deployed within the risk assessment process. In this same perspective, Jajagandan[43] et al. outline the general steps to consider in a risk analysis, steps that are fully aligned with ISO 15189 requirements. They emphasize the implementation and follow-up of actions. To go further, P. Pernas, in his article dedicated to the management of corrective actions[44]—whether they follow a risk analysis or a non-conformity—underlines the importance of their effective implementation to avoid the recurrence of the same problem. He also demonstrates the dynamic nature of risk analysis, contrary to the static aspect often attributed to it. Thus, risk analysis alone cannot be sufficient. Within a continuous improvement approach in the laboratory, it must be accompanied by the rigorous implementation of the actions decided upon in response to an identified risk, the monitoring of their effectiveness, and continuous surveillance.

Following the risk analysis, a relocation plan was established. In a facility with a quality management system, as described in Chapter 5 of the ISO 15189:2022 standard, an organizational hierarchy is defined, and each stakeholder has specific responsibilities. Within the biochemical genetics laboratory, activities were placed under the supervision of the laboratory manager, the quality manager, and the scientists in charge of BabyDetect analysis. Since the relocation had been carried out, activities also fell indirectly under the responsibility of those in charge of the analyses performed at the Unilab facilities. As BabyDetect analysis involves multiple areas of expertise, activities were also indirectly placed under the supervision of the bioinformatics team. In this context, all proposals relating to the validation plan had to

be reviewed and validated by all the teams concerned, each within its respective areas of expertise.

In such a multidisciplinary environment, it is not always easy to reach a consensus regarding the implementation modalities of the relocation, the management of the associated process, or the organization of activities in the new premises after transfer. The quality manager, also playing a mediating role among the different team members, must maintain a neutral position and carry out a critical and objective analysis of the various proposals. This approach aims to ensure that the laboratory's activities continue to run smoothly. Furthermore, this type of situation also represents an opportunity to identify and propose areas for organizational and managerial improvement, particularly with a view to possible similar future situations.

As mentioned above, and under the responsibility of the different teams, a validation plan was established before the relocation. This plan was modified after the transfer, resulting in the content presented in the section devoted to the validation plan in the results section. The modifications made consisted mainly of additions to the initial plan, particularly in the second part of the relocation concerning the validation of library preparation and sequencing.

This situation highlights one of the fundamental principles of quality management, namely the Deming cycle: Plan-Do-Check-Act. In a manner of speaking, within the principle of improvement, a plan can absolutely be designed and then modified with the aim of being improved. It is also a matter of defining the activities to be carried out and implementing what has been defined. Although the additional elements had not been foreseen in the initial validation plan, it was essential, from a quality perspective, to integrate the new data into the validation file, and to ensure its documentation as well as its traceability, in accordance with the requirements of paragraphs 7.3 and 8.4 of the ISO 15189:2022 standard, relating respectively to the validation of analytical methods and the management of records. Paragraph 7.3 specifies that the laboratory must validate examination methods after significant modification and that records relating to the validations carried out, including the results obtained, must be retained. Paragraph 8.4 stipulates that records must be created for each activity that impacts the quality of an examination at the time it is performed. Reports of results must remain accessible for as long as necessary or required, be legible regardless of the storage medium used by the laboratory and remain available for laboratory management reviews.

Overall, the results obtained after the relocation confirm that the BabyDetect test can continue to be offered, as the analyses have demonstrated that the results it provides remain reliable after the transfer of activities. Thus, the relocation did not have a negative impact on the quality of the results generated by the BabyDetect test.

2. Phase 1: Revalidation of the extraction step

It should be highlighted that, within the framework of the BabyDetect analysis performed after the relocation, the size measurement of the DNA fragments extracted using the Nucleocube 1 automated system was carried out. This measurement was performed using the Fragment Analyzer available within the GIGA, a CHU unit collaborating with the genetic biochemistry laboratory, and more specifically with the BabyDetect activity. This unit also provides the sequencers used in the context of the present validation/revalidation.

However, unlike the sequencers, the Fragment Analyzer is not listed in the laboratory's quality management system. In this context, the applicable requirements of Annex XI of the IVDR regulation as well as those of chapter 6.4.7 of the ISO 15189:2022 standard are not met. These requirements notably mandate the maintenance of records for any equipment capable of influencing the results of laboratory activities, including, among others: the demonstration of equipment conformity with defined acceptance criteria, the preventive maintenance program, the location of the instrument, as well as the information associated with its management.

It is therefore necessary to ensure the conformity of the equipment used, the availability and control of the associated records, as well as their integration and traceability within the quality management system.

3. Phase 2: Revalidation of the Remaining Targeted NGS Panel Workflow

Despite the observed statistically significant differences for certain parameters, the analytical performance of the workflow was preserved. No exceeding of the defined quality thresholds was observed after the relocation, and for certain parameters and samples, performance was even better after the transfer.

Indeed, a statistical difference is not necessarily correlated with an analytical difference. The results show that, despite the observed statistical differences, the quality parameters remain well above the defined acceptability thresholds. This reflects the maintenance of the overall performance of the workflow as well as its robustness against the observed variations in concentration. Furthermore, the inter-run comparisons carried out before the relocation, as well as the results obtained when comparing the same samples before and after the relocation, support the hypothesis that the observed variability is mainly linked to the inter-run effect rather than a real effect of the relocation. Moreover, the absence of a significant difference in post-capture concentrations reflects the good standardization of the workflow prior to the sequencing step.

In the context of NGS, inter-run variations regarding coverage, duplication rate, or other quality metrics are naturally expected, even under stable analytical conditions. Following a study[45] on

the reproducibility of variant calling, variability in quality metrics between replicate samples within the same run was reported, which was also observed in our study when comparing the quality metrics of sample QI-14-00BD00-011601 sequenced in intra-run duplicate 116. Furthermore, pre-analytical conditions remain inherently variable, notably due to pipetting, initial DNA quality, storage conditions, or quantification steps.

The observed differences may also be influenced by the sensitivity of statistical tests, which increases with sample size. Indeed, as the number of samples increases, the standard error decreases and estimations become more precise, which can lead to the detection of statistically significant differences even when these remain analytically small. This is particularly the case when comparing the means obtained before and after relocation for the quality parameters "% Read aligned" and "%Q30". On this subject, Greenland[46] et al. describe the impact of sample size on statistical power as well as the limitations of an interpretation based solely on p-values in decision-making.

Moreover, a statistical difference does not necessarily translate into a clinically relevant difference. In the present case, the same variants were identified after relocation for the samples whose results were already known prior to the transfer. Likewise, internal controls remained negative in both situations. The absence of an impact on variant detection constitutes the fundamental objective of the validation, which aims to demonstrate the method's ability to correctly detect variants while indirectly attesting to analytical reproducibility, as well as the maintenance of proficiency and robustness of the bioinformatics pipeline used for variant analysis. This aspect was confirmed by the concordance observed between the VCF files as well as the variants detected before and after relocation.

4. Validation of the Franklin Interpretation Software

Differences were observed regarding the positions of certain variants for specific genes at the transcript level. This was notably the case for the *G6PD* gene (sample 14-248048-0020), where we found: c.949G>A in Franklin and c.1039G>A in Alissa. These differences have no impact on the type of variant identified, the molecular interpretation and the associated protein.

A difference in transcripts was also observed for other genes without impacting the position of the detected variant, notably for the *BTD* gene (sample 14-248038-0054), for which we obtained the variant c.1270G>C with Alissa and c.1270G>C with Franklin. For all cases, see Appendix BGE.BABYDETECT.ANA.A01d in the Vivaldi document control software.

Furthermore, the BED files sent to Twist and the decision tree (Tri-tree) built in Alissa were identically transferred to the Franklin platform. The bioinformatics team also identified a difference in the transcripts between the Humanomics BED file and the Franklin BED file,

specifically for the *KH1*, *MACF1*, *CACNA1D*, *CACNA1C*, *GNAS*, *PLEC*, *POR*, and *WNK1* genes. This was corrected by modifying the BED transcripts in Humanomics to match those in Franklin. This shows the importance of involving all teams in the revalidation, in this case the bioinformatics team.

5. NGS-WES Validation

With a view to expanding the number of diseases that can be screened early in children, but also to improve the flexibility of the test, a screening approach based on the filtered exome was considered within the framework of the BabyDetect test. This approach would notably allow the identification of certain rare diseases associated with copy number variations, while improving the investigation of diseases already screened for by the targeted panel thanks to the expansion of the analyzed regions. The flexibility offered by the exome, particularly regarding the addition or removal of regions of interest, also constitutes a major advantage from a validation perspective, as it would avoid a complete revalidation of the method each time a new pathology of interest is added pathologies that are currently not covered by the BabyDetect test.

However, the question of the scope of genetic screening remains central, particularly in order to maintain a balance between clinical benefit and the risk of overinterpretation. As such, it appears essential to prioritize not only the clinical relevance of the diseases or genes investigated, but also the associated therapeutic possibilities. Beyond strict medical considerations, the expansion of screening also raises important ethical and societal questions. This issue therefore requires a multidisciplinary approach involving different stakeholders, or at the very least, consideration of their respective points of view. These include physicians, parents, scientists involved in test development, and policy makers.

From this perspective, a study[47] conducted in the Netherlands examined the views of different stakeholders regarding the expansion of the national newborn screening program. Regarding the expansion of screening to well-characterized and curable diseases, all stakeholders were in favor of this expansion, provided that the initial objective of the program remained unchanged. In contrast, opinions diverged regarding incidental findings of non-curable diseases. Some parents and healthcare professionals believed it was useful to communicate this information when known, despite the absence of available treatment. Conversely, some patients preferred not to be informed, while some practitioners considered that such information could influence parents in their decisions regarding possible future pregnancies. Thus, the expansion of newborn screening presents significant challenges that must be carefully assessed and controlled before any large-scale implementation.

V. CONCLUSION

The BabyDetect test is an analysis enabling the screening of rare genetic diseases (1/2000 births) at birth. It utilizes an innovative next-generation sequencing approach with the aim of highlighting early on treatable genetic diseases chosen by considering well-defined inclusion criteria. Considering the regulatory context in which we were operating, notably under the IVDR regulation and the ISO 15189-2022 standard, the objective was to meet their requirements following major changes that occurred in the performance of the BabyDetect test.

These changes were: the relocation of the activity from the initial premises (GIGA Tower, +3) to the new Unilab-Lg premises (Tower 6, +1); the change of variant analysis and interpretation software (from Alissa to Franklin); and the evolution of the screening method, transitioning from a targeted NGS-panel to a filtered NGS-exome. Within this framework, it was necessary to conduct a revalidation and a method validation to demonstrate the maintenance of both analytical and bioinformatic performances following these various changes.

The risk analysis was conducted, highlighting various critical elements to focus on. It allowed for the implementation of means to control these critical points to guarantee the proper continuity of the analysis after the relocation. These means notably involved planning through action plans that were created and executed. Furthermore, these actions led to satisfactory results: the revalidation of the targeted NGS-panel method demonstrated that the relocation had no impact on analytical variability, analytical performance, and even less so on bioinformatic performance.

The validation of the Franklin software demonstrated its capacity to identify and interpret variants in a reliable and reproducible manner, thereby guaranteeing the continuity of the BabyDetect analysis following the shutdown of the previous Alissa interpretation platform. Finally, the first tests performed within the framework of the filtered NGS-exome method allowed us to verify that the biochemistry designed and provided by Twist for library preparation works. This will enable proceeding with the validation step on a larger number of samples.

Beyond the analytical aspect, this work highlighted the role of quality assurance in a laboratory. More precisely, the management of major changes within an ISO 15189 and IVDR context. The importance of mastering the life cycle of analytical methods was brought to light. The modifications, whether technical or related to documentation, required, firstly, evaluating the associated risks, their probability of occurrence, and especially their impact on the analysis. Secondly, a realization or implementation step was necessary, leading to results that, each time, require being recorded and traceable. This aligns directly with the famous Deming Plan-Do-Check-Act cycle and data integrity principles. The outcome of this cycle is the assurance

of maintaining quality and therefore the reliability of the results produced by the laboratory. This approach thus illustrates the central role of a quality management system in the laboratory.

NGS analysis in general, although innovative, still harbors certain limitations that we previously enumerated in section 6.2 of our introduction. Continuous monitoring regarding variant interpretation and classification strategies is necessary in view of the constant evolution of knowledge and technologies/software in medical genetics.

The results obtained from the clinical trial phase of the BabyDetect test up to the present offer perspectives to consider for the evolution of neonatal screening. On one hand, they underline the benefit of reflecting on expanding access to the test, the benefit of which currently remains limited by the lack of reimbursement by the INAMI. On the other hand, they suggest that the developed approach could contribute, in the long term, to the evolution of the national neonatal screening program by enabling the early identification of an increased number of rare genetic diseases by meeting the established criteria for public health screening.

With the rapid and growing evolution of sequencing technologies, also considering the various modifications within the European regulatory framework, it remains vital to maintain a quality approach to guarantee the compliance and reliability of laboratory results, but even more so the safety and health of patients.

BIBLIOGRAPHY

1. Clinical Biology C. *Practical guideline for the implementation of quality system in accredited clinical biology laboratories version 4-2025*. 2025.
2. Public Health and Environment E. *Royal Decree of December 14, 1987 setting the standards that human genetics centers must meet (amended by the RD of January 25, 1989)*. 1988.
3. ISO 5649:2023. *Medical laboratories — Concepts and specifications for the design, development, implementation, and use of laboratory-developed tests*. https://www.iso.org/obp/ui/fr/#iso:std:iso:5649:ed-1:v1:fr (accessed 18 Mar 2026).
4. Lasserre A, Naour N, Zagury P. *Newborn Screening: Evaluation Criteria for the Inclusion of New Diseases in the National Newborn Screening Program*. Saint-Denis La Plaine, France: French National Authority for Health (HAS); 2023. ISBN 978-2-11-167588-9.
5. Woolf L, Adams J. The early history of PKU. *Int J Neonatal Screen*. 2020;6(3):59.
6. Cheillan D. Main biological tools applied to newborn screening – Current status and future perspectives. *médecine/sciences*. 2021;37:461–467.
7. Loeber J, et al. Newborn screening in Europe: Evolution over the last decade and analysis of the current situation by the International Society for Neonatal Screening. *médecine/sciences*. 2021;37:441–456.
8. Wilson J, Jungner G. *Principles and Practice of Screening for Disease*. Geneva: World Health Organization; 1968. Public Health Papers No. 3

9. Neonatal screening. Why screen? https://www.depistageneonatal.be/depistage-danomalies-congenitales/pourquoi-depister/ (accessed 23 Mar 2026).

10. Government of the French Community of Belgium. Decree of 9 January 2020 on the screening of congenital anomalies in the French Community. Belgian Official Gazette (Moniteur belge), 7 February 2020, No. 2020010313.

11. BOUCHOUX G., SABLIER M. Mass spectrometry – Principle and equipment. Techniques de l'Ingénieur, dossier P2645, 2005.

12. Boemer F, et al. A next-generation newborn screening pilot study: NGS on dried blood spots detects causal mutations in patients with inherited metabolic diseases. *Sci Rep*. 2017;7:17641.

13. Gutierrez-Mateo C, et al. Development of a multiplex real-time PCR assay for the newborn screening of SCID, SMA, and XLA. *Int J Neonatal Screen*. 2019;5(4):39. doi:10.3390/ijns5040039.

14. Boemer F, et al. *May the sun arise on SMA: Hope for a spinal muscular atrophy-free region*. [No further details provided].

15. Hovhannesyan K, et al. Analytical validation of a genomic newborn screening workflow. *Int J Neonatal Screen*. 2025;11(4):91. doi:10.3390/ijns11040091.

16. Lunke S, et al. Feasibility, acceptability and clinical outcomes of the BabyScreen+ genomic newborn screening study. *Nat Med*. 2025;31:4236–4245.

17. Wang X, et al. Combined genetic screening and traditional biochemical screening to optimize newborn screening systems. *Clin Chim Acta*. 2022;528:44–51.

18. Stroek K, et al. Evaluation of 11 years of newborn screening for maple syrup urine disease in the Netherlands and a systematic review of the literature: Strategies for optimization. *JIMD Rep*. 2020;54:68–78.

19. Khoja I, et al. Reduction of false-positive results with biochemical second-tier testing for newborn screening of Pompe disease. *Genet Med*. 2026;28:101604.

20. Malvagia S, et al. Development of strategies to decrease false positive results in newborn screening. *Int J Neonatal Screen*. 2020;6(4):84. doi:10.3390/ijns6040084.

21. Yu B, Yang Y, Zhou L, Wang Q. Evaluating a novel newborn screening methodology: Combined genetic and biochemical screenings. *Arch Med Res*. 2024;55:102959.

22. Adam M, et al. GeneReviews glossary. In: *GeneReviews® [Internet]*. University of Washington, Seattle; 2026. https://www.ncbi.nlm.nih.gov/books/NBK5191/ (accessed 24 Mar 2026).

23. Slatko B, Gardner A, Ausubel F. Overview of next-generation sequencing technologies. *Curr Protoc Mol Biol*. 2018. doi:10.1002/cpmb.59.

24. Shah P, Nathanson K. Application of panel-based tests for inherited risk of cancer. *Annu Rev Genomics Hum Genet*. 2017;18:201–227.

25. Racu M, et al. Validation of a targeted next-generation sequencing panel for pancreatic ductal adenocarcinomas. *Exp Mol Pathol*. 2024;139:104920.

26. Shubina J, et al. WES-based screening of 7,000 newborns: A pilot study in Russia. *Hum Genet Genomics Adv*. 2024;5:100334.

27. Bros-Facer V, Taylor S, Patch C. Next-generation sequencing-based newborn screening initiatives in Europe: An overview. *Rare Dis Orphan Drugs J*. 2023;2:N/A.

28. Rehm HL, et al. ACMG clinical laboratory standards for next-generation sequencing. *Genet Med*. 2013;15:733–747.

29. Aziz N, et al. College of American Pathologists' laboratory standards for next-generation sequencing clinical tests. *Arch Pathol Lab Med*. 2015;139:481–493.

30. Kiewiet G, et al. Future of Dutch NGS-based newborn screening: Exploring the technical possibilities and assessment of a variant classification strategy. *Int J Neonatal Screen*. 2024;10:20.

31. Miller DT, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2021 update: A policy statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2021;23:1391–1398.

32. ISO 15189:2022. *Medical laboratories — Requirements for quality and competence*. ISO. https://www.iso.org/fr/standard/76677.html (accessed 3 Apr 2026).

33. MDCG 2022-6. *Guidance on significant changes regarding the transitional provision under Article 110(3) of the IVDR*. Public Health. https://health.ec.europa.eu/latest-updates/mdcg-2022-6-guidance-significant-changes-regarding-transitional-provision-under-article-1103-ivdr-2022-05-04_en (accessed 25 May 2026).

34. Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on in vitro diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU. 2025. http://data.europa.eu/eli/reg/2017/746/2025-01-10 (accessed 31 Mar 2026).

35. Devices manufactured and used only within health institutions ('in-house devices'). *AFMPS*. https://www.afmps.be/fr/usage_humain/produits_de_sante/dispositifs_medicaux_et_leurs_accessoires/etablissements_et_2 (accessed 3 Apr 2026).

36. MDCG endorsed documents and other guidance. *Public Health*. 2026. https://health.ec.europa.eu/medical-devices-sector/new-regulations/guidance-mdcg-endorsed-documents-and-other-guidance_en (accessed 3 Apr 2026).

37. COFRAC. *Technical accreditation guide for verification (scope A) / validation (scope B) of medical biology methods*. SH GTA 04 - Révision 02 Available online:www.cofrac.fr.

38. COFRAC. *Technical accreditation guide for high-throughput sequencing (NGS) technology*. SH GTA 16 - Révision 00. Available online:www.cofrac.fr.

39. Illumina MD. *DRAGEN secondary analysis: Calling of precise, efficient, and comprehensive variants for next-generation sequencing data*. Technical sheet.M-GL-00680 FRA v13.0.

40. Olson N, et al. Precision FDA Truth Challenge V2: Calling variants from short and long reads in difficult-to-map regions. *Cell Genomics*. 2022;2:100129.

41. Behera S, et al. Comprehensive genome analysis and variant detection at scale using DRAGEN. *Nat Biotechnol*. 2025;43:1177–1191.

42. O'Rawe J *et al.* Low concordance of multiple variant-calling pipelines: practical implications for exome and genome sequencing. *Genome Med* 2013; **5**: 28.

43. Jayamani J, Janardan CC, Appan SV, Kathamuthu K, Ahmed ME. A practical tool for risk management in clinical laboratories. *Cureus*. 2022;14:e32774.

44. Pernas P, members of the Quality Management subgroup. Quality improvement: Recommendations for risks and irregularities management. *Ann Biol Clin (Paris)*. 2013;71 Spec No 1:95–113.

45. Qi Y, et al. Reproducibility of variant calls in replicate next generation sequencing experiments. *PLoS ONE*. 2015;10:e0119230.

46. Greenland S, et al. Statistical tests, P values, confidence intervals, and power: A guide to misinterpretations. *Eur J Epidemiol*. 2016;31:337–350.

47. van Dijk T, et al. Expanding neonatal bloodspot screening: A multi-stakeholder perspective. *Front Pediatr*. 2021;9:706394. doi:10.3389/fped.2021.706394.

48. Federal Agency for Medicines and Health Products (FAMHP/AFMPS). Questions and Answers on In-house Devices under IVDR. Brussels; latest version consulted [Jun 2026].