

## Qualitätssicherung im Rahmen der Genomdiagnostik

Florian Kraft<sup>1\*</sup>, Anna Benet-Pagès<sup>2,3\*</sup>, Daniel Berner<sup>4</sup>, Anna Teubert<sup>5</sup>, Sebastian Eck<sup>6</sup>,  
Norbert Arnold<sup>7</sup>, Peter Bauer<sup>8</sup>, Matthias Begemann<sup>1</sup>, Marc Sturm<sup>9</sup>, Stephanie Kleinle<sup>3</sup>,  
Tobias Haack<sup>9</sup>, Thomas Eggermann<sup>1</sup>

<sup>1</sup>Institute für Humangenetik und Genommedizin, Medizinische Fakultät der RWTH Aachen,  
Aachen, DE

<sup>2</sup>Institut für Neurogenomik, Helmholtz Zentrum München, Neuherberg, DE

<sup>3</sup>Medizinisch Genetisches Zentrum, München, DE

<sup>4</sup>MVZ genetikum GmbH, Neu-Ulm, DE

<sup>5</sup>amedes genetics, Hannover, DE

<sup>6</sup>Bionano, San Diego, CA 92121, USA

<sup>7</sup>Zentrum für familiären Brust- und Eierstockkrebs; Klinik für Gynäkologie und Geburtshilfe,  
Universitätsklinikum Schleswig-Holstein, DE

<sup>8</sup>CENTOGENE N.V., Rostock, DE

<sup>9</sup>Institut für Medizinische Genetik und Angewandte Genomik, Universität Tübingen,  
Tübingen, DE

\*These authors contributed equally.

### Korrespondenzadresse:

Prof. Dr. rer. Nat. T. Eggermann  
Institut für Humangenetik und Genommedizin  
Pauwelsstr. 30, D-52074 Aachen  
Tel.: 0241 8037285  
Fax: 0241 8082580  
Email: [teggermann@ukaachen.de](mailto:teggermann@ukaachen.de)

**Dieses Manuskript wird in englischer Sprache in der Medizinischen Genetik 2023  
veröffentlicht.**

### **Hinweis:**

*In diesem Dokument stellt ein Team von Autor\*Innen mit Erfahrungen in den verschiedenen Bereichen der NGS-basierten humangenetischen Diagnostik – vom Nasslabor über die Datenverarbeitung und Befundung bis zur Qualitätssicherung – Aspekte vor, die im Rahmen der Qualitätssicherung der Genomdiagnostik berücksichtigt werden sollten. Es erhebt keinen Anspruch auf Vollständigkeit und die Autor\*innen übernehmen keinerlei Haftung. Unabhängig von den folgenden Ausführungen müssen die Leser\*Innen die konkreten Sachverhalte gemäß den aktuellen Entwicklungen und Gegebenheiten sorgfältig und eigenständig prüfen und auf dieser Grundlage eine Entscheidung über die ggf. notwendige Anpassung des hausinternen Vorgehens und dessen Dokumentation treffen. Neben den Organisationen, die Richtlinien z.T. auf nationaler Ebene organisieren bzw. vorschlagen (Gendiagnostikkommission GEKO, Deutsche Akkreditierungsstelle DakkS, Bundesärztekammer BÄK, Deutsche Gesellschaft für Humangenetik GfH e.V.) sei auch auf internationale Gremien verwiesen, die Dokumente und Vorschläge zur Qualitätssicherung im Bereich der Genomdiagnostik formulieren, z. B. die Global Alliance for Genomics and Health (GA4GH), die Europäische Gesellschaft für Humangenetik (ESHG: EuroGentest) und die Medical Genome Initiative.*

### **Zusammenfassung**

Die rasante und dynamische Implementierung von Next Generation Sequencing (NGS)-basierenden Ansätzen hat die genetische Testung wesentlich verändert. In naher Zukunft werden nahezu alle molekularen Veränderungen im menschlichen Genom durch Hochdurchsatzsequenzierung analysierbar sein. Dieser Fortschritt führt zu einer zunehmend zentralen Rolle des Faches Humangenetik im multidisziplinären Management von Patienten mit genetischen Veränderungen. Es muss aber auch von qualitätssichernden Maßnahmen begleitet werden, die den sicheren und optimalen Einsatz des erworbenen Wissens in der Genomdiagnostik gewährleistet. Um dieses Ziel zu erreichen, wurden zahlreiche Instrumentarien und Leitlinien entwickelt. In diesem Manuskript fassen Autor\*innen mit verschiedenen Expertisen im Bereich Genomanalytik den derzeitigen Status der Qualitätssicherung in der Genomdiagnostik zusammen mit dem Ziel, eine weitere Standardisierung und Qualitätsverbesserung des NGS als Kernkompetenz unseres Faches zu erzielen.

## 1. Einleitung

Der Einsatz von Next-Generation Sequencing (NGS)-basierten Analysen in der humangenetischen Routinediagnostik hat die Nachweisrate von klinisch relevanten genomischen Veränderungen bei Patienten mit verschiedenen Krankheitsbildern wesentlich erhöht und erlaubt eine umfassende, kosten- und zeiteffiziente Abklärung genetisch heterogener Erkrankungen. Daher sind NGS-basierte Befunde zunehmend auch die Grundlage für personalisierte Therapieansätze. Die hohe Relevanz einer schnellen genetischen Diagnostik für das Management der Patienten (und ihrer Familien) wird am Beispiel der Versorgung akut kranker Kinder deutlich (z.B. (1)). Damit ist die humangenetische Diagnostik zu einer zentralen, fächerübergreifenden Disziplin geworden, bei der die humangenetische Kernkompetenz der Varianteninterpretation in der multidisziplinären Patientenbetreuung unerlässlich ist.

Die Dynamik der NGS-basierten Analytik äußert sich neben dem o.g. zunehmenden Bedarf an genetischer Testung in der rasanten technologischen Entwicklung sowohl im Nasslabor als auch in der bioinformatischen Prozessierung und anschließenden Auswertung von NGS-Daten. Wurden zu Beginn des routinediagnostischen NGS-Einsatzes primär Multigenpanel eingesetzt, um indikationsspezifische Analysen zu ermöglichen, so ist zumindest als technische Grundlage einer Auswertung der zunehmende Einsatz von Exom-Sequenzierungen (ES) zu beobachten. Die erzeugten ES-Daten werden in Abhängigkeit von der klinischen Fragestellung ausgewertet, entweder weiterhin genspezifisch in Hinblick auf krankheitsursächliche Veränderungen für ausgesuchte Phänotypen oder exomweit bei klinisch heterogenen bzw. unspezifischen Krankheitsbildern. Mit beiden Anreicherungs- und Auswertestrategien können Einzelbasenveränderungen und Insertions-Deletions-Varianten (SNVs/InDels) sowie Kopienzahlabweichungen (CNVs) detektiert werden. Andere Variantentypen wie Repeatexpansionen (REs) oder weitere strukturelle Varianten wie Translokationen und Inversionen (im Folgenden SV genannt) entgehen aber methodisch bedingt zum großen Teil der Detektion und müssen derzeit mit anderen Methoden gezielt untersucht werden. Auch ist die Abdeckung der zu analysierenden Zielregion nicht homogen und erfordert eine entsprechende Berücksichtigung bei der Auswertung und Befundung (Typ A-C Tests (2)). Pseudogene und Transkriptvarianten können ebenfalls ein Auswerteproblem darstellen. Internationale Projekte mit dem Ziel einer konvergierenden Annotation von Transkripten wie MANE (*Matched Annotation from the NCBI and EMBL-EBI*), die die Definition eines

genomweiten Satzes von Referenz-Transkripten erarbeiten, sind hier notwendige Initiativen zur Lösung dieser Herausforderungen (3).

Die (Weiter)Entwicklung und diagnostische Implementierung von methodischen Ansätzen wie die Genom-Sequenzierung (GS) (zur Übersicht: (4)) tragen zur Verbesserung für verschiedene Typen pathogenetisch relevanter Veränderungen bei, z.T. ergänzt durch Verfahren wie z.B. Long-read Sequencing und Transkriptom-Analysen, bei. Der Mehrwert der GS gegenüber der ES ergibt sich neben einer Implementierung der bioinformatischen Algorithmen zur Detektion zusätzlicher Varianten-Typen insbesondere aus der Ausweitung des Untersuchungsumfangs über die Zielregion der proteinkodierenden Exons hinaus. Aufgrund derzeit noch unzureichender Frequenzdaten und Annotation von Varianten in nicht-kodierenden Bereichen stellt die GS für diagnostische Labore eine Herausforderung dar.

Die Komplexität der NGS-Analytik inklusive der methodischen Limitationen und Dynamik, aber auch der stetig wachsende Bedarf an humangenetischer Diagnostik bei gleichzeitigem Zeitdruck erfordert die Formulierung von Qualitätsanforderungen. Diese sollte ein Labor im Rahmen der NGS-Diagnostik erfüllen, um zum maximalen Benefit der Patienten eine moderne und qualitativ hochwertige auf internationalem Standard anbieten zu können.

Im Folgenden werden in Anlehnung an bereits publizierte internationale Leitlinien und Empfehlungen zur Qualitätssicherung Vorschläge vorgestellt, die zur standardisierten qualitätsgesicherten NGS-Diagnostik beitragen sollen.

Die Autor\*innen beziehen sich insbesondere auf die von EuroGentest formulierten Empfehlungen, die in Grundlagen bereits 2016 formuliert und zum großen Teil von der S1-Leitlinie „Molekulargenetische Diagnostik mit Hochdurchsatz-Verfahren der Keimbahn, beispielsweise mit Next-Generation Sequencing (2018)“ übernommen wurden, und die kürzlich im Hinblick auf die diagnostische Nutzung des GS erweitert wurden (2, 5, 6). Die Ausführungen sind dabei sowohl für die konstitutionelle als auch somatische NGS-Diagnostik im Bereich Humangenetik anwendbar.

Die Gliederung der folgenden Ausführungen richtet sich nach der derzeitigen Gliederung der DIN EN ISO 15189 (im Folgenden: ISO15189) als formale Richtschnur der Qualitätssicherung, und ordnet dieser – soweit möglich – die o.g. Empfehlungen (recommendations) von EuroGentest („EuroGen recom.“ (5)) zu. Ergänzend berücksichtigt werden dabei weitere Vorgaben wie die Richtlinien der Bundesärztekammer (s. Link: RiliBÄK) und weitere relevante Beiträge zur Thematik (z.B. (7, 8)). Neben den bereits genannten Dokumenten hinsichtlich der Qualitätssicherung für NGS Analysen gibt es weitere Empfehlungen, z.B. von der Global Alliance for Genomics and Health (GA4GH), der Deutschen Gesellschaft für

Huangenetik (GfH), der Gendiagnostikkommission (GEKO) oder auch der Medical Genome Initiative.

## **2. Allgemeine Überlegungen**

Exom- und genomweite Analysen sollten dann eingesetzt werden, wenn eine relevante Verbesserung der Qualität, Effizienz und Detektionsrate für die Abklärung der pathogenetischen Ursachen eines Krankheitsbildes zu erwarten sind (EuroGen recom. 1 (5)). Bei spezifischen genetischen Erkrankungen mit definierter genetischer Ursache können aber auch gezielte genetische Untersuchungen am Beginn der Diagnostik ausreichend sein. Eine stufenweise Vorgehensweise sollte dem spezifischen klinischen Auftrag entsprechend - wenn immer möglich - angemessen sein (EuroGen recom. 5 (5)). Da das klinische Bild bei vielen genetisch bedingten Erkrankungen aber stark variieren und überlappen kann, ist eine gezielte molekulare Testung oftmals schwierig. NGS-basierte Diagnostikansätze bieten daher die Möglichkeit, dieses breite Spektrum von molekularen Veränderungen in einem Testansatz zu erfassen. Trotzdem müssen die Angemessenheit des Tests und Bearbeitungszeit (z.B. bei Therapierelevanz) bei der Beauftragung berücksichtigt werden.

Ferner sollte das Labor den Umgang mit Zusatz- bzw. Zufallsbefunden (9) im Vorfeld der Implementierung insbesondere von exom- und genomweiten Verfahren festlegen (EuroGen recom. 31) und ggf. den Einsendern kommunizieren, auch vor dem Hintergrund der Qualifikationsanforderungen an die beauftragende ärztliche Person (10). Vorschläge verschiedener Fachgremien sollten hier Berücksichtigung finden (z.B. ClinGen Actionability working group, American College of Medical Genetics (11-13)).

### **2.1 Gesetzliche und normative Vorgaben zu Qualitätssicherung**

In Bezug auf rechtliche und normative Vorgaben unterscheidet sich die Qualitätssicherung im Rahmen der Genomdiagnostik nicht von anderen Verfahren der humangenetischen Diagnostik. Allerdings ergeben sich aufgrund des Umfangs der anfallenden Daten und der o.g. möglichen Zusatz- bzw. Zufallsbefunde sowie der möglichen Nutzung von Geräten und Softwarelösungen außerhalb des direkten Überwachungsbereichs der Einrichtung ein zusätzlicher Regelungsbedarf im Rahmen der gesetzlichen und normativen Vorgaben.

Grundsätzlich gelten die Anforderung des Gendiagnostikgesetzes (GenDG; <http://www.gesetze-im-internet.de/gendg/index.html>). Die Umsetzung der Diagnostik hat sich an der RiliBÄK zu orientieren. Die in Deutschland für die genetische Diagnostik nicht

zwingend vorgeschriebene Akkreditierung nach ISO15189 gibt hier einen zusätzlichen Rahmen vor. In diesem Zusammenhang ist darauf hinzuweisen, dass zwar derzeit die Revision der ISO15189 von 2012 gilt, dass die Version 2022 aber nach einer Übergangsfrist ab 2025 umzusetzen ist. Weitere Normen, die je nach Tätigkeitsschwerpunkt von Laboren genutzt werden, sind die DIN EN ISO17020 (anwendbar für den Bereich Pathologie/Neuropathologie) und DIN EN ISO17025 (Prüf- und Kalibrierlabore). Im Bereich Softwareentwicklung kann die IEC 62304 (für „Medizingeräte-Software“) für Anforderungen an und Entwicklung von Software in der genetischen Diagnostik Hilfestellung leisten. Seit 2022 ist zusätzlich zu diesen meist gezielt auf genetische Analysen abzielenden Gesetzen und Normen auch die Verordnung (EU) 2017/746 über In-vitro-Diagnostika umzusetzen.

## **2.2 Zielsetzung der Genomanalyse**

Diagnostische Genomanalysen werden mit dem Ziel durchgeführt, sowohl hereditäre als auch somatische Ursachen von Erkrankungen zu finden. Dies bedeutet, dass die Analyse unterschiedlicher Primärproben, wie Blut und Gewebe, abhängig von der gestellten Indikation notwendig sein können. In der Diagnostik finden derzeit vor allem „Short-read“-Sequenzieretechniken mit Leselängen von 100-300 bp Anwendung, die in Abhängigkeit von den genetischen Besonderheiten der zu analysierenden Gene durch „Long-read“-Sequenzierungen mit Leselängen von 10-100 kb komplementiert werden können(14).

Bei einer Reihe syndromaler Erkrankungen werden pathogene Varianten regelhaft im Mosaikstatus nachgewiesen (z.B. Proteus-Syndrom (AKT1), Tuberöse Sklerose (TSC1)). Bei Erkrankungen wie der Neurofibromatose Typ 1 können Varianten in betroffenem Gewebe im Mosaik vorliegen, im Blut aber möglicherweise nicht nachweisbar sein. Darüber hinaus findet der Nachweis somatischer Varianten Anwendung als Biomarker für das Therapieansprechen onkologischer Patienten, z.B. bei der Therapie mit PARP-Inhibitoren (15). Es sollten daher an die genetische Fragestellung angepasste Bioinformatik-Pipelines und Qualitätsparameter mit einer ausreichenden Nachweisgrenze (Limit of Detection, LoD) verwendet werden (5). Zu einer umfassenden humangenetischen Abklärung kann auch der Nachweis von Varianten des mitochondrialen Genoms gehören, mit Berücksichtigung von Besonderheiten dieser Varianten wie „Mosaikstatus“ in unterschiedlichen Geweben (16, 17). Zur Validierung gehört neben der Bestimmung der Sensitivität und Präzision die Erhebung der LoD für die Detektion heteroplasmischer mitochondrialer Varianten. Das Vorhandensein von NUMTs (nuclear

mitochondrial DNA sequences, „mitochondriale Pseudogene“) im Kerngenom kann zu Artefakten bei der Erkennung und Interpretation von Varianten führen.

Bioinformatische Tools ermöglichen neben der Erkennung von SNVs, CNVs und SVs aber auch die Detektion von „Runs of homozygosity-ROH“, die Hinweise auf das Vorliegen einer Uniparentalen (Iso-)Disomie (UPD) und damit einer autosomal-rezessiven oder Imprinting-Erkrankung geben können. In Trio-Analysen sind UPD auch ohne ROH-Detektion nachweisbar. Auch sekundäre, somatische Ereignisse im Sinne einer klonalen Hämatopoese, die zu einem kopienneutralen Loss-of-Heterozygosity führen und ggf. auch pathogene Veränderungen maskieren können, können über ROH-Analysen erfasst werden (18).

Neben der Diagnostik monogener, oligogener und (Tumor-)Prädispositionen werden Verfahren zum Nachweis von Varianten auch für die Bewertung multifaktorieller Erkrankungen und pharmakogenetischer Fragestellungen verwendet. Die Berechnung von polygenen Risiko-Scores zur Bestimmung relativer Risiken auf der Grundlage von Assoziationsstudien hat inzwischen ebenfalls Einzug in die humangenetische Diagnostik gehalten.

Bei der Durchführung pränataler Analysen spielt der Ausschluss einer maternalen Kontamination weiterhin eine zentrale Rolle. Dies kann im Rahmen einer pränatalen Trio-Diagnostik direkt aus den NGS-Daten möglich sein. Die Analyse zellfreier fetaler DNA im mütterlichen Plasma zum Nachweis von Aneuploidien und weiteren genomischen Veränderungen im Rahmen der nicht-invasiven pränatalen Diagnostik (NIPT) ist mit spezifischen Qualitätsanforderungen verbunden (19).

### **3. Aspekte der Qualitätssicherung im Rahmen der NGS-basierten Diagnostik**

Die im Folgenden beschriebenen Aspekte sollten in Anlehnung an die ISO15189 im Hinblick auf qualitätssichernde Maßnahmen im Rahmen der NGS-basierenden Diagnostik berücksichtigt werden. Auch sind die Inhalte angelehnt an die EuroGentest-Empfehlungen (5), die fordern, dass Labore, die NGS-Analysen durchführen, akkreditiert sein müssen (EuroGen recom. 2). Diese Empfehlung ist in Deutschland im GenDG nicht verpflichtend vorgesehen, allerdings müssen Einrichtungen, die genetische Analysen zu medizinischen Zwecken durchführen, gemäß §5(2) die Anforderungen erfüllen, die in §5(1) für die akkreditierungspflichtigen Untersuchungen zur Klärung der Abstammung genannt sind.

### **3.1 Personal**

Das in die Diagnostik eingebundene Personal muss in Abhängigkeit von der Aufgabe (Präanalytik, Probenerfassung, Nasslabor, Bioinformatik, Befundung, Validierung, Medizinische Bewertung) qualifiziert sein und über die angemessene Ausbildung, Erfahrungen und Fertigkeiten verfügen (s. ISO15189:2012: 5.1). In diesem Zusammenhang ist neben den gesetzlichen Vorgaben durch das GenDG auch die S2k-Leitlinie der deutschen Gesellschaft für Humangenetik (GfH) (20) zu beachten, die neben einer personell ausreichenden Ausstattung für die verantwortliche Supervision der technischen Abläufe, Dateninterpretation und Befunderstellung die Einbindung entsprechend ausgebildeter Naturwissenschaftler\*innen (z.B. Fachhumangenetiker\*innen) und Ärzt\*innen fordert. Arbeitsplatzbezogene Einarbeitungen und fortwährende Fortbildungen sind durchzuführen und zu dokumentieren. Für die in Datenbewertung und -befundung verantwortlich eingebundenen Personen ist neben der beruflichen Qualifikation (s.o.) eine mehrjährige diagnostische Erfahrung in der molekulargenetischen Analytik von Sequenzveränderungen für genetisch heterogene Erkrankungen (inkl. der Auswertung und Variantenbewertung) erforderlich.

### **3.2 Räumlichkeiten**

Die Räumlichkeiten müssen bezüglich Ausstattung, Platz, Aufteilung und Zustand (Elektrik, Klimatechnik, Statik, Zugangskontrolle) geeignet sein, die Anforderungen einer modernen NGS-Diagnostik zu erfüllen. So sind kreuzkontaminationsgefährdete Arbeitsschritte zu trennen. Entsprechend ist die getrennte Lagerung von Primärproben, Nukleinsäuren, Kontrollmaterialien, Reagenzien, Verbrauchsmaterialien und Analyseprodukten zu beachten. Im Bereich Datenverarbeitung ist in Hinblick auf die Hardware-Komponenten eine räumliche Unterbringung vorzusehen, die eine physikalische Sicherung der Daten gewährleistet.

### **3.3 Verbrauchsmittel, Geräte und Software**

Das Laboratorium muss ein dokumentiertes Verfahren zur Beschaffung, Handhabung inkl. Instandhaltung, Wartung, Lagerung und Absetzung der Ausrüstung verfügen. Chargen von Verbrauchsmaterialien müssen normkonform erfasst und geführt werden. Für Geräte muss eine (Funktions-)Überwachung gemäß ISO15189 dokumentiert werden. Dies bezieht alle in die Prozesse einbezogenen Geräte ein, vom Kleingerät (z.B. Pipetten) über Thermocycler und Messgeräte zur Qualitäts- und Quantitätsbestimmung bis zu den NGS-Plattformen und



zugehörigen EDV-Systemen. Die notwendige Software muss ebenfalls im Hinblick auf die relevanten Module beschrieben sein. Bei Implementierung müssen alle Komponenten und Module der Prozessabläufe in den Validierungsmaßnahmen berücksichtigt sein und die Validierung ggf. modular dokumentiert werden. Maßgebliche Änderungen inkl. Neuversionierungen der Nasslaborchemie, der Sequenziergeräte oder der Auswertesoftware müssen neu verifiziert bzw. validiert werden.

Aufgrund der hohen Investitions- bzw. Wartungskosten für NGS-Plattformen und -Auswertelösungen werden diese z.T. von verschiedenen Einrichtungen gemeinsam genutzt, bzw. in Hinblick auf Analysesoftware und Datenspeicherung wird auf externe, u.a. kommerzielle, Anbieter zurückgegriffen. Diese Kooperationen sind prinzipiell möglich, aber die einschlägigen Bestimmungen (Datenschutzgrundverordnung, GenDG, ISO15189) müssen berücksichtigt. So muss eine gemeinsame Gerätenutzung dergestalt dokumentiert und geregelt sein, dass die mit der genetischen Untersuchung beauftragte Einrichtung eine vollständige Kontrolle über den Ablauf behält und dass keine unbefugte Person Zugriff auf das Gerät und damit die Korrektheit und Sicherheit der Patientendaten und -proben erhält. Der Einsatz von Auswertungslösungen, die cloudbasiert genutzt werden, muss neben den o.g. Sicherheiten in Bezug auf Datenschutz und probenbezogene eindeutige Identifizierbarkeit berücksichtigen, dass das Labor jederzeit auf die Daten zugreifen kann und Verfügbarkeit und Speicherung über den Zeitraum möglich ist, für den eine Dokumentationspflicht gesetzlich vorgeschrieben ist.

### **3.4 Präanalytik**

Die für NGS-Analysen notwendigen präanalytischen Maßnahmen müssen neben den von der ISO15189 beschriebenen Anforderungen (Anweisung zur Probenentnahme, -lagerung) insbesondere die von §8 des GenDG geforderte Einwilligung zur Probenentnahme, Durchführung der Untersuchung sowie Verarbeitung der genetischen Daten vorhalten. Dabei sind die Vorgaben zur Aufklärung der GEKO inkl. der Qualifikationsanforderungen an die beauftragende ärztliche Person zu berücksichtigen (10).

Wie in den ESHG-Leitlinien zur NGS-Diagnostik (2) und in den S1-Leitlinien der GfH (6) formuliert, ist es die Aufgabe des Labors, den Einsendern eine öffentlich zugängliche Liste mit krankheitsrelevanten Genen („Zielgene“ im angebotenen Panel) zur Verfügung zu stellen, die den Untersuchungsumfang definiert. Auch müssen Angaben zur Art der genetischen Variante (SNVs, CNVs, SV, Repeats), für die der Test validiert ist, sowie zu Limitationen der NGS-Diagnostik gemacht werden (EuroGen recom. 5, 6 (5)).

### 3.5 Durchführung von NGS-Verfahren

Im Folgenden werden die verschiedenen Arbeitsschritte der diagnostischen NGS-Abläufe unter besonderer Berücksichtigung qualitätssichernden Maßnahmen vorgestellt (Abb. 1).

#### Übersicht über die wesentlichen Elemente der Qualitätssicherung im Rahmen von diagnostischen NGS-Strategien

Testentwicklung	Testvalidierung	QM/Produktüberwachung
<ul style="list-style-type: none"><li>• Definition des Testzwecks (z.B. WES)</li><li>• Identifizierung bestgeeigneter Wetlab- und Auswertestrategien</li><li>• Definition von Qualitätsmerkmalen</li><li>• Etablierung des Tests/der Teststrecke</li><li>• Risikoabschätzung der einzelnen Prozessschritte</li></ul>	<ul style="list-style-type: none"><li>• Bestimmung der Leistungsmerkmale und Effizienz</li><li>• Bestimmung der Reproduzierbarkeit und Limitationen</li><li>• Prüfung, Verwendung und Dokumentation von Referenzmaterialien und -sequenzen</li></ul>	<ul style="list-style-type: none"><li>• Arbeitstäglige Überprüfung der Qualitätsmerkmale</li><li>• Regelmäßige Sequenzierung von Kontrollproben mit bekanntem Genotyp</li><li>• Regelmäßige Teilnahme an Ringversuchen (indikations- und methodenspezifisch, RiLiBÄK-konform)</li></ul>

Abbildung 1: Formale Ebenen der Qualitätsanforderungen im NGS-Prozess.

#### 3.5.1 Nasslabor

Wie für andere molekulargenetische Arbeitsabläufe ist grundsätzlich auf eine räumliche und gerätetechnische Trennung von Arbeitsbereichen zur Vermeidung von Kontaminationen zu achten. Aufgrund der im Vergleich zu konventionellen molekulargenetischen Verfahren zum Teil komplexeren Abläufen mit mehreren PCR-basierten Schritten, die die o.g. Trennung sehr erschweren, muss bei Abweichung von einer strikten Trennung vor Ort geprüft und ggf. eine Risikobewertung vorgenommen werden. So kann die in Prozessabläufen notwendige Dokumentation von Qualitätsparametern (s.u.), aber auch die Prüfung der Auswertedaten im Hinblick auf möglich Kontaminationen eine Abschätzung des Risikos ermöglichen. Bei einzelnen, nicht-PCR basierten Verfahren erübrigen sich diese Überlegungen gegebenenfalls. Für den Nasslaborprozess müssen alle verwendeten Verfahren und deren Schritte durch entsprechende Arbeitsanweisungen beschrieben sein. Dabei müssen neben dem eigentlichen Arbeitsablauf auch die Dokumentation von analyserelevanten Chargen und qualitätsrelevanten Daten der NGS-Anreicherung festgelegt sein und die zugehörigen Parameter müssen definiert

sein. Dies gilt auch für die Auswerteparameter und deren Historie (s.u.). Erfolgt die Dokumentation der Daten EDV-basiert, müssen die entsprechenden Module im Hinblick auf Dokumentationslenkung, Datensicherheit und Validierung beschrieben sein.

Labortechnische und organisatorische Maßnahmen zur eindeutigen Identifizierung und Rückverfolgung der Proben müssen implementiert werden (ESHG: statement 16 (2)).

### 3.5.2 Datenprozessierung

Eine NGS-Bioinformatik-Pipeline umfasst Algorithmen, die Sequenzierdaten und zugehörige Metadaten verarbeiten, von den Rohdaten bis zur Varianten-Annotation und Varianten-Priorisierung. Bei diesem Prozess können mehrere Softwarekomponenten, Datenbanken und Betriebsumgebungen (Hardware und Betriebssystem) zum Einsatz kommen. Außerdem werden

**Tabelle 1:** Short-read next generation sequencing Datei-Formate. (\*BAM und CRAM Dateien enthalten auch nicht alignierte Sequenzen.)

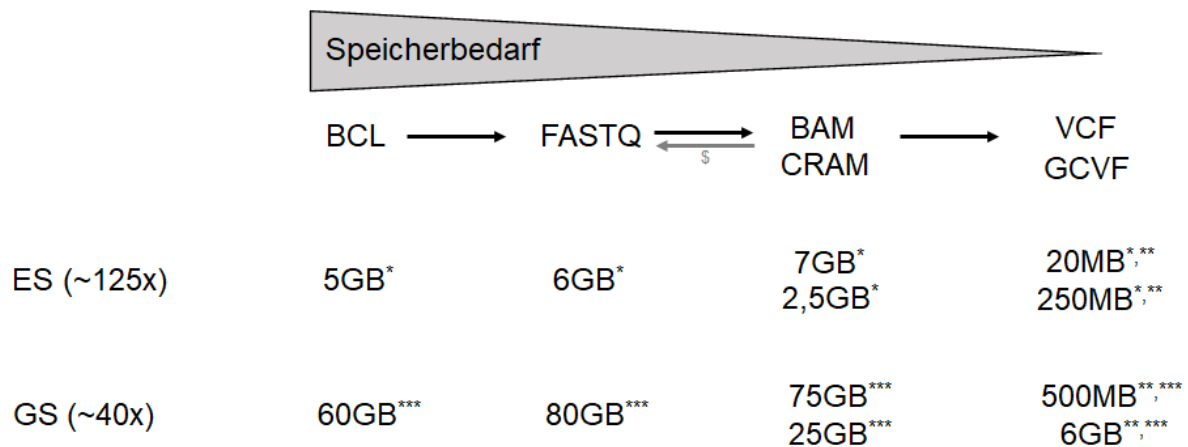
Dateityp	Inhalt
BCL	binary base call) – Enthält die Sequenzrohdaten
FASTQ	Sequenzierte Reads mit korrespondierenden Qualitäts-Werten.
BED	Browser Extensible Data: Generisches Format zur Angabe von genomischen Regionen als Kombination von genomischen Koordinaten und Annotationen.
SAM	<i>Sequence Alignment Map</i> : Informationen über Reads und deren Mapping im Genom.
BAM*	<i>Binary Alignment Map</i> : Komprimierte binäre Version des SAM files.
CRAM*	<i>Compressed Reference-oriented Alignment Map</i> : Stark komprimierte Version des SAM files.
VCF	<i>Variant Call Format</i> : Generisches Format zur Beschreibung von genetischen Varianten.
GVCF	<i>Genomic VCF</i> Format mit Informationen bezüglich jeder Base im Genom unabhängig vom Vorliegen einer Variante

bei der Datenanalyse verschiedene Dateitypen verwendet und generiert (Tab. 1, Abb. 2), Standard-Datenformate wie FASTQ, BED, SAM, BAM, CRAM und VCF sind weit verbreitet und werden auch empfohlen (EuroGen recom. 17).

Der Arbeitsablauf für klinische NGS-basierende Datenprozessierung umfasst drei wesentliche Prozessschritte: Die Primäranalyse zur Generierung von DNA-Sequenzdaten aus den Rohdaten, die Sekundäranalyse, die u.a. das Read-Alignment zu einem Referenzgenom und

Varianten-Calling umfasst, und die Tertiäranalyse zur Varianten-Annotation, Filterung, Interpretation und Berichterstattung (7). Obwohl sich der bioinformatische Prozess für die Analyse von GS nicht wesentlich von dem für ES unterscheidet, sind oft zusätzliche

Algorithmen z.B. für die Erfassung von SV, Repeatexpansionen und Annotation von nicht-kodierenden Bereichen notwendig.



**Abbildung 2:** Speicherbedarf im Rahmen der NGS-Diagnostik am Beispiel Exom- und Genom-Sequenzierung. Dargestellt sind der Datenfluss und die beispielhafte Größe von ES- bzw. GS-Datensätzen. (\* Durchschnittliche Dateigröße für 1 Sample bei 125x Abdeckung und einer Readlänge von 2x150 Basen. Die Größe der Dateien kann je nach eingesetzter Geräteplattform (aufgrund der verschiedenen Kodierung des Basen-Qualitätsscores), Enrichment und Kompressionslevel leicht variieren. \*\* Durchschnittliche Größe für VCF-Dateien, welche mittels „GATK best practise workflow“ generiert wurde. Diese kann je nach verwendetem Variantencaller, Filterstrategie und Kompressionslevel deutlich variieren. \*\*\*Durchschnittliche Dateigröße für 1 Sample bei 40x Abdeckung und einer Readlänge von 2x150 Basen. Die Größe der Dateien kann je nach verwendetem Sequenzierer (aufgrund der verschiedenen Kodierung des Basen-Qualitätsscores) und Kompressionslevel leicht variieren. <sup>§</sup>Die Rückkonvertierung von BAM/CRAM in FASTQ ist nicht verlustfrei, da in BAM/CRAM-Dateien nur der erste Teil des Illumina Read Namens übernommen werden. Die Information bezüglich Reads 1 oder 2 bei paired-end Sequenzierung, Pass Filter und der Indexsequenz gehen verloren.)

### 3.5.2.1 Primäranalyse

Die Primäranalyse umfasst das Basecalling, also die Generierung von Sequenz- und Qualitätsdaten (Phred Scores), und Demultiplexen der Rohdaten welche i.d.R. im FASTQ-Format gespeichert werden. Dabei ist der Phred Score ein wichtiges Maß für die Richtigkeit des individuellen Calls einer Base und kann z.B. durch technische Artefakte beeinflusst werden. Hier ermöglicht u.a. der „GATK best practise workflow“ durch die Rekalibrierung des Basisqualitätsscores (BQSR) über maschinelles Lernen die teilweise Korrektur dieser Fehler. Artefakte in NGS-Alignments sind geräteunabhängig und führen häufig zu falsch-positiven De-novo-Varianten. Bei der Variantenfilterung können systematische Fehler durch den Vergleich mit Datensätzen, die auf demselben Sequenziergerät erzeugt wurden, reduziert werden. Durch Entfernen von Adaptoren und Trimmen von Basen mit schlechter Qualität wird sichergestellt, dass nur qualitativ hochwertige Reads mit der optimalen Länge für die nachgeschaltete Analyse verwendet werden.

### 3.5.2.2 Sekundäranalyse

Die Sekundäranalyse ist der rechenintensivste Teil einer NGS-Analyse. Sowohl die Wahl des Referenzgenoms (GRCh37/hg19, GRCh38/hg38, T2T-CHM13) als auch des Aligners (z.B. für DNA BWA-MEM, DRAGMAP) können große Auswirkungen auf die Genauigkeit komplexer Varianten-Calls haben (21). Eine verbesserte Variantendetektion und Genotypisierung wurde für das kürzlich veröffentlichten T2T-CHM13 Assembly nachgewiesen (22), dessen Anwendung für die klinische Diagnostik allerdings noch nicht ausreichend bewertet ist. Eine breite diagnostische Anwendung scheitert z.Z. noch an der fehlenden Übertragung von Varianten-Datenbanken (z.B. ClinVar, gnomAD) auf die genomischen Positionen im T2T-CHM13 Genom. Benchmarking-Studien zeigen, dass die Verwendung von GRCh38/hg38 im Vergleich zu GRCh37/hg19 die Variantendetektion verbessert und weniger falsch-positive Varianten generiert. Die Verwendung wird daher von Eurogentest empfohlen (EuroGen recom. 15 (5)). Allerdings sind teilweise Datensätze in Datenbanken nur für GRCh37 bzw. GRCh38 verfügbar, sodass die „vollständige“ Annotation von Varianten dann nur mit einem liftover von Daten möglich ist. Aufgrund von fälschlicherweise im GRCh38-Assembly integrierten Duplikationen, ist die Variantendetektion für mehrere Gene (z.B. *CBS*, *KCNE1* and *CRYAA*) im Vergleich zu GRCh37 schlechter. Dieses Problem kann aber durch die Maskierung dieser Bereiche im Referenzgenom behoben werden (21). Ein Ansatz zur Verbesserung der Variantendetektion in diesen Genen und weiteren mit homologen Kopien im Genom (z.B. *SMN1*, *PKD1*, *NEB*, *GBA1*, *STRC*) ist die Verwendung von Sequenzgraphen in den Alignment-Algorithmen (23).

Alignierte Sequenzen werden in komprimierten BAM- oder CRAM-Dateien zusammen mit der Positionsinformation der Sequenz im Genom hinterlegt (Tab. 1, Abb. 2). Zur Vermeidung von Artefakten und falsch-positiven Varianten sollten Duplikate markiert bzw. entfernt und ein lokales Re-Alignment vorgenommen werden. Im Gegensatz zur ES werden die Library-Präparationen beim GS, sofern ausreichend DNA vorhanden ist, PCR-frei generiert. Dies reduziert den Anteil duplizierter Reads auf 3-5 %, im Vergleich zu 5-15 % in einem typischen ES (24). Zum anderen werden GC-reiche Bereiche besser erfasst, was zu einer einheitlichen Sequenziertiefe und diagnostischen Abdeckung führt.

Wie beschrieben eröffnet der Einsatz des GS die Möglichkeit, komplexe Variantentypen zu erkennen, die mit ES nur bedingt zu erfassen sind (SVs, REs). Es wird empfohlen, dass Pipelines verschiedene Variantendetektions-Tools nutzen (Tab. 2), die für jede Klasse von Varianten spezifisch sind, um die Sensitivität der Detektion zu maximieren (24). Die

Kombination der Ergebnisse von zwei orthogonalen SNV-/InDel-Callern (z. B. GATK HaplotypeCaller, DeepVariant, Strelka2) und die anschließende Zusammenführung mehrerer Varianten-Callsets (im VCF-Format) zu einem einzigen unter Verwendung von Softwarepaketen wie BCFtools kann einen leichten Sensitivitäts-Vorteil bieten. Allerdings sollte darauf geachtet werden, dass komplexe Varianten und/oder Unterschiede in der Variantendarstellung korrekt behandelt werden (8).

**Tabelle 2:** Häufig eingesetzte Software-Module für Variant-Calling, Quality-Check und andere notwendige Prozessierungsschritte beim Short-read-Sequencing. (Angaben zur genannten Software können gesondert zur Verfügung gestellt werden.)

Analyseschritt	Prozessierungsschritt	Software (beispielhaft)
Primary analysis	Demultiplexing	bcl2fastq
	Alignment	BWA-Mem, DRAGMAP
	File management	BCFtools
	SNV calling	GATK HaplotypeCaller, DeepVariant, Strelka2
Secondary analysis	CNV calling	MOPS, CNVkit, CONTRA, CoNVEX, ExomeCNV, ExomeDepth, GATK, XHMM
	Structural Variants calling	DELLY, Lumpy, Manta, Pindel, SVMerge
	Repeat expansions	ExpansionHunter, exSTRa, STRetch, TREDPARSE
	TEs and insertions	Mobster, MELT, TraFiC-mem
	Low frequency variants (mosaic and somatic)	MuTect2

Das Labor sollte sowohl den Einfluss des Aligners als auch die Genauigkeit der Variantenerkennung für alle nachweisbaren Variantenklassen validieren, d.h. für Einzelnukleotidvarianten (SNVs) und InDels, Kopienzahlvarianten (CNVs) und komplexe genomische Strukturvarianten (SVs, REs) (EuroGen recom. 22 (5)). Die Sequenzierung des gesamten Genoms bietet den umfassendsten Ansatz und liefert in der Regel eine durchschnittliche Sequenztiefe von etwa 30-60x über das gesamte Genom. Die höhere Sequenztiefe, die bei der Panel- und Exom-Sequenzierung erreicht wird, kann jedoch eine sensitivere Erkennung von Varianten mit geringer Allelfrequenz ermöglichen, z. B. subklonale somatische Mutationen bei Tumoren oder Varianten, die als Keimbahnmosaik vorliegen (24). Dabei sind auch die Zuverlässigkeit und Genauigkeit von Algorithmen zur Detektion von SVs und REs aus GS-Daten z.T. noch nicht optimal. Sie erlauben aber die Detektion einer Vielzahl

von pathogenen Veränderungen und werden derzeit eher als Screening-Verfahren eingesetzt. Vor der Mitteilung solcher komplexer Varianten ist ggf. deren technische Bestätigung mit einer alternativen diagnostisch etablierten Methode erforderlich.

### 3.5.2.3 Tertiäranalyse

Die Variantenannotation als erster Schritt der Tertiäranalyse benennt jede Variante anhand einer standardisierten Nomenklatur und verknüpft sie mit Informationen aus verschiedenen Datenbanken, der medizinischen Literatur und mit Informationen über die Qualität der Variantendetektion (z. B. die Anzahl der Reads, in denen die Variante auftritt). Die Nomenklaturrichtlinien für SNVs der Human Genome Variation Society (HGVS) und für CNVs des International Standing Committee on Human Cytogenetic Nomenclature (ISCN) sollten hier genutzt werden (Tab. 3).

**Tabelle 3:** Weiterführende Angaben zu Nomenklaturrichtlinien und Referenzmaterialien (Auswahl).

Name	Website
Human Genome Variation Society (HGVS)	<a href="https://varnomen.hgvs.org/">https://varnomen.hgvs.org/</a>
Human Gene Nomenclature (HUGO)	<a href="https://www.genenames.org/">https://www.genenames.org/</a>
Matched Annotation from the NCBI and EMBL-EBI (MANE)	<a href="https://www.ncbi.nlm.nih.gov/refseq/MANE/">https://www.ncbi.nlm.nih.gov/refseq/MANE/</a>
International Society of Cytogenetic Nomenclature (ISCN)	<a href="https://iscn.karger.com/">https://iscn.karger.com/</a>
Genome-in-a-bottle	<a href="https://www.nist.gov/programs-projects/genome-bottle">https://www.nist.gov/programs-projects/genome-bottle</a>

Die korrekte Annotation von Varianten ist für die anschließende Filterung, Priorisierung, Klassifizierung und Interpretation von Varianten erforderlich (s. 3.6). Die Annotation ist bei komplexen genomischen Varianten immer noch eine Herausforderung, wie z.B. im Fall von MNVs (multinucleotide variants), die zwar meist unabhängig voneinander aufgelistet werden, aber gemeinsam betrachtet werden müssen, um ihre Pathogenität richtig bewerten zu können. Auch die korrekte Detektion und Annotation von komplexen strukturellen Variationen wie Inversionen und balancierte Translokationen kann mit GS-Daten alleine erschwert sein und den Einsatz einer weiteren Methode erforderlich machen.

Es wird empfohlen, alle GS-Varianten eines Standortes zu annotieren (EuroGen recom. 23) und die aggregierten Daten zu kompilieren, um interne Allelhäufigkeiten zu generieren und die interpretierten Daten mit anderen Einrichtungen zu teilen (EuroGen recom. 24)(5).

### 3.5.2.4 Qualitätskontrolle der Pipeline und Validierungs-Zyklen

Die Bioinformatik-Pipeline sollte über eine ausreichende Qualitätskontrolle verfügen, um sicherzustellen, dass die erzeugten Daten robust sind und Fehlererkennung sowie Einhaltung der Vorgaben gegeben sind (EuroGen recom. 16, 21 (5)). Zur Qualitätskontrolle des Bioinformatik-Prozesses können in jedem Analyseschritt Qualitätsdaten verwendet werden (Tab. 4). Die NGS-Pipeline sollte plattformspezifisch und auf die Bedürfnisse des Labors zugeschnitten sein (EuroGen recom. 17 (5)).

**Tabelle 4:** Relevante Qualitätsparameter im NGS-Ablauf, die u.a. auch im Rahmen der EMQN-Ringversuche erfasst werden.

Parameter	Bedeutung
Yield	Durchsatz des Sequenzierlaufs.
Error Rate	Prozentsatz der falsch bestimmten Basen pro Zyklus, bestimmt durch Alignment an PhiX.
%Q30	Prozentsatz der Basen, die mit einem Qualitätswert von 30 (entspricht 1 Fehler in 1000 Basen, Phred-Score) oder mehr bestimmt wurden.
Density (K/mm <sup>2</sup> )	Cluster Dichte in tausend pro mm <sup>2</sup> .
Cluster PF (%)	Cluster Dichten, die den internen Chastity Filter genügen und damit für die Downstream Analyse genutzt werden können.
Phas / Prephas (%)	Prozentsatz an Molekülen eines Clusters, die in jedem Zyklus außer Sync geraten durch Einbau multipler Nukleotide (Pre-Phasing) oder durch fehlgeschlagenen Einbau eines Nukleotids (Phasing).
Percentage mapped reads	Prozentsatz an Reads, der an die Referenzsequenz gemappt werden konnten.
Duplicates (%)	Prozentsatz an Duplikaten, die normalerweise nicht für Downstream-Analysen genutzt werden.
Insert Size	Durchschnittliche Größe der sequenzierten DNA-Moleküle.
GC content (%)	Durchschnittlicher GC Gehalt der sequenzierten DNA-Fragmente.
Mean Coverage	Durchschnittliche Anzahl von Reads, die eine bestimmte Zielregion abdecken.
Cov30	Die durchschnittliche Anzahl der sequenzierten Basen, die mit bekannten Referenzbasen übereinstimmen oder diese "abdecken". Ein ganzes Genom, das mit einer 30-fachen Abdeckung sequenziert wurde, bedeutet beispielsweise, dass jede Base im Genom im Durchschnitt 30 Mal sequenziert wurde.
Uniformity	Coverage-Uniformität beschreibt die Verteilung der Reads über die Zielregionen im Genom.



Bioinformatik-Pipelines, die für die Analyse klinischer GS/ES-Daten entwickelt wurden, sind komplex und erfordern ein robustes Qualitätssicherungsprogramm für die laufende Überwachung und die Aktualisierung der Pipelines. Validierungen müssen alle Schritte dokumentieren, die zum Kompilieren, Installieren und Ausführen spezieller klinischer GS/ES-Bioinformatik-Pipelines erforderlich sind. Ein Versionskontrollsystem und eine ausführliche Dokumentation von Code Änderungen sind für jede Bereitstellung von Software-Updates erforderlich ((25); IEC 62304).

Pipelines sollten anhand von Referenzstandards (z.B. Genome-in-a-bottle Proben) getestet werden, um sicherzustellen, dass sie reproduzierbar und fehlerfrei funktionieren (26). Die Validierung erfolgt an einer bestimmten Version der Software, und sollte bei jeder Aktualisierung von Softwarekomponenten und Referenzdatendateien wiederholt werden. Die Laboratorien sollten über ein Release-Management mit einer klaren Definition von "größeren" und "kleineren" Änderungen verfügen, die den Grad der Validierung definieren. Wenn Softwarekomponenten oder Parameter maßgeblich geändert werden, sollte der Testmechanismus die lokale Testleistung und die Auswirkungen dieser Änderungen auf den klinischen Prozess der Variantendetektion und der Annotation erneut validieren. Idealerweise sollte die Softwarevalidierung kontinuierliche Integrationsprozesse für Aktualisierungen und Erweiterungen umfassen. Dies kann es erforderlich machen, dass die Softwaretests bei inkrementellen Änderungen bis zu einem gewissen Grad automatisiert werden. Es ist von entscheidender Bedeutung, dass Fehler wirksam aufgespürt und zugrunde liegende Probleme identifiziert und behoben werden.

### **3.5.3 NGS-Qualitätsparameter**

Wichtige Qualitätsparameter (Tab. 4) für die NGS-Analytik können für jeden Lauf über Tools wie fastQC, multiQC oder Illumina Sequence Analysis Viewer (SAV) erfasst werden. Bei der Auswertung ist darauf zu achten, möglichst die Metriken pro Read und pro Lane zu betrachten, da zum Beispiel ein Abfall der Qualität in einem bestimmten Bereich der Flowcell in einer globalen Run-Statistik übersehen werden kann.

Je nach Methode und Anwendung können sich Akzeptanzgrenzen für einzelne Werte unterscheiden, z.B. ist für ES-Anwendungen eine durchschnittliche Coverage der Zielregionen von 100-120X anzustreben, während bei einer GS schon 30-40X ausreichend sein können. Unter dem Gesichtspunkt der diagnostischen Qualität ist des Weiteren die Abdeckung der Zielregionen mit einer definierten Mindestabdeckung (z.B. 20X), die einen sicheren Nachweis

oder Ausschluss einer Veränderung erlaubt, entscheidend. Bei der Zielregion sollte dabei zwischen der Zielregion zur Beurteilung des Assays (z.B. Zielregion der Anreicherung) und der diagnostisch relevanten Zielregion (z.B. Consensus Coding Sequence (CCDS) plus/minus 10 bp) unterschieden werden. In PCR-freien GS-Analysen werden hierbei üblicherweise für die Autosomen Werte von >99% mit mindestens 20-facher Abdeckung erreicht. Im Kontext der Fragestellung ist je nach Anwendung zu bewerten, wie sehr ein Abweichen von den gesetzten Qualitäts-Cutoffs die Qualität des Untersuchungsergebnisses beeinflusst.

Bei spezifischen Fragestellungen, z.B. in Hinblick auf Mosaikauswertung und im Tumorbereich müssen ggf. weitere Überlegungen zur Qualitätssicherung wie Sequenziertiefe, berichtete Allelfrequenzen und tumorspezifische Variantenbewertung angestellt werden.

### **3.6 Postanalytik**

In diesem von der ISO15189 definierten Bereich wird u.a. die Anforderungen an die Freigabe von Untersuchungsergebnissen formal erfasst. Die Postanalytik sollte neben den bereits beschriebenen Kriterien zur Ergebnisbewertung auch die Plausibilitätskontrolle beinhalten, z.B. um bei nicht hinreichend sicheren Untersuchungsergebnissen im Hinblick auf den Untersuchungsauftrag mit einer alternativen Methode zusätzliche Informationen zur Sicherung zu erheben (z.B. Sanger-Sequenzierung, MLPA).

#### **3.6.1 Variantenbewertung**

Die Variantenbewertung ist eine der zentralen Kernkompetenzen des Faches Humangenetik, da sie die Grundlage für die weitere Beratung und (personalisierte) Therapie für Patienten mit genetisch verursachten Erkrankungen darstellt. Aufgrund der zentralen Bedeutung der Variantenbewertung für die genetische Beratung, Prädiktion und/oder Therapie, muss die Bewertung der nachgewiesenen genetischen Veränderungen nachvollziehbar nach einem einheitlichen Standardverfahren gewährleistet sein. Die Variantenbewertung sollte in Anlehnung an das 2015 ausgearbeitete fünf Klassen System des „American College of Medical Genetics and Genomics (ACMG) / Association for Molecular Pathology (AMP) erfolgen (27): benigne, wahrscheinlich benigne, Variante unklarer Signifikanz (VUS), wahrscheinlich pathogen und pathogen. Viele Labore verwenden in Anlehnung an die ACMG-Einteilung eine Einstufung in die Klassen 1 (benigne) bis 5 (pathogen), allerdings sollte hierbei beachtet werden, dass diese Klassifizierung nicht international genutzt wird und ggf. zu Missverständnissen führen kann. Wird eine solche Klassifizierung verwendet, muss eine

Erläuterung der Klassifizierung in Anlehnung an die ACMG-Vorschläge u.a. im Befund gegeben werden.

In der ACMG-Einteilung werden insgesamt 16 pathogene und 12 benigne Kriterien zur Klassifizierung herangezogen, die sich in ihrer jeweiligen Evidenzstärke unterscheiden. Auf Basis dieser fünf Klassen haben sich in den letzten Jahren eine Vielzahl an Empfehlungen zur Anwendung und Auslegung der einzelnen ACMG/AMP-Kriterien entwickelt. Hervorzuheben ist hierbei die Empfehlung zur Variantenklassifikation der „Association for Clinical Genomic Science“ (ACGS)(s. Links), die unter anderem nicht nur Vorschläge ausspricht, wie die einzelnen Kriterien verwendet und kombiniert werden können, sondern auch eine mögliche Hoch- bzw. Abstufung der Evidenzstärke für die jeweiligen Kriterien angibt. Diese Empfehlungen sollten für eine korrekte Variantenbewertung mit einbezogen werden. Ebenfalls zu berücksichtigen sind eine Reihe von wichtigen generellen Empfehlungen in Bezug auf die Verwendung der Evidenzkriterien, die die „Sequence Variant Interpretation Working Group“ (SVI WG) von „Clinical Genome Resource“ (ClinGen, s. Links) ausgearbeitet hat. Hier wird für einige Kriterien zudem ein punktebasiertes System vorgeschlagen, dass als Entscheidungshilfe zur Bewertung der einzelnen Evidenzstärken mit herangezogen werden sollte.

Die Verwendung geeigneter Datenbanken ist für eine korrekte Variantenbewertung unumgänglich. Neben den allgemeinen Varianten-, Populations- und Sequenzdatenbanken, die größtenteils bereits in der ACMG/AMP-Publikation von 2015 aufgeführt sind, gibt es auch eine Reihe von weiteren Gen- bzw. Phänotyp-spezifischen Datenbanken, die für spezifische Indikationen herangezogen werden sollten (Tab. 5). Zudem sollten die Gen-spezifischen Handreichungen der „Variant Curation Expert Panels“ (VCEP) von ClinGen bei der Variantenbewertung entsprechender Gene berücksichtigt werden, bei der explizit die Auslegung und Anwendung der einzelnen Evidenzkriterien für das jeweilige Gen von einem Fachgremium ausgearbeitet wurden. Eine Übersicht hierzu findet sich im „ClinGen Criteria Specification“ (CSpec) Registry (<https://cspec.genome.network/cspec/ui/svi/>). Die Einschätzungen der VCEP sowie die Bewertungen einiger Varianten aus den Gen- bzw. Phänotyp-spezifischen Datenbanken findet sich mittlerweile auch in den generellen Datenbanken wie ClinVar oder LOVD wieder (Tab. 5).

Datenbanken wie die o.g. ClinVar und LOVD, aber auch die neu aufgestellte HGQN (Human Genetics Quality Network, Datenbank des BVDH), werden nicht nur für die Bewertung von Varianten im eigenen Labor genutzt, sondern dienen auch dazu, neu nachgewiesene Varianten

öffentlich zugänglich zu machen. Auf diese Weise erfolgt ein Erkenntnisaustausch über mögliche Pathogenitätseinstufungen dieser neuen, aber auch bereits bekannter Varianten.

**Tabelle 5:** Auswahl an Datenbanken, die allgemein zur Variantenbewertung eingesetzt werden, aber auch Daten zu bestimmten Genen enthalten.

<b>Variantedatenbanken</b>	<b>Inhalt bzw. spezifisch erfasste Gene und Phänotypen</b>	<b>Website</b>
ClinVar	Genomweit	<a href="https://www.ncbi.nlm.nih.gov/clinvar/">https://www.ncbi.nlm.nih.gov/clinvar/</a>
LOVD	Genomweit	<a href="https://www.lovd.nl/">https://www.lovd.nl/</a>
HGQN	Genomweit	<a href="https://www.hgqn.org/">https://www.hgqn.org/</a>
DECIPHER	Genomweit	<a href="https://www.deciphergenomics.org/">https://www.deciphergenomics.org/</a>
BRCA Exchange (ENIGMA)	<i>BRCA1, BRCA2</i>	<a href="https://brcaexchange.org/?=">https://brcaexchange.org/?=</a>
CanVar	Tumorprädispositionssyndrome	<a href="https://canvaruk.org/">https://canvaruk.org/</a>
CFTR1 / CFTR2	<i>CFTR</i>	<a href="http://www.genet.sickkids.on.ca/Home.html">http://www.genet.sickkids.on.ca/Home.html</a> <a href="https://cftr2.org/resources">https://cftr2.org/resources</a>
DVD	Hörstörungen	<a href="https://deafnessvariationdatabase.org/">https://deafnessvariationdatabase.org/</a>
Emhg	<i>RYR1, CACNA1S</i>	<a href="https://www.emhg.org/diagnostic-mutations">https://www.emhg.org/diagnostic-mutations</a>
Infervers	Autoinflammatorische Erkrankungen	<a href="https://infervers.umai-montpellier.fr/web/index.php">https://infervers.umai-montpellier.fr/web/index.php</a>
InSiGHT	<i>APC, CDH1, EPCA, GALNT12, MLH1, MSH2, MSH6, MUTYH, PMS2</i>	<a href="http://insight-database.org/">http://insight-database.org/</a>
MITOMAP	Mitochondriale Erkrankungen	<a href="https://www.mitomap.org/MITOMAP">https://www.mitomap.org/MITOMAP</a>
PKD	<i>PKD1, PKD2</i>	<a href="https://pkdb.mayo.edu/variants">https://pkdb.mayo.edu/variants</a>
RettBASE	<i>MECP2, CDKL5, FOXP1</i>	<a href="http://mecp2.chw.edu.au/">http://mecp2.chw.edu.au/</a>
<b>Populationsdatenbanken</b>		
gnomAD	genomweit	<a href="https://gnomad.broadinstitute.org/">https://gnomad.broadinstitute.org/</a>
dbSNP	Genomweit	<a href="https://www.ncbi.nlm.nih.gov/SNP/">https://www.ncbi.nlm.nih.gov/SNP/</a>
DGV	Apathogene Strukturvarianten	<a href="http://dgv.tcag.ca/dgv/app/home">http://dgv.tcag.ca/dgv/app/home</a>
MitoPhen	Patienten mit pathogener mtDNA-Veränderung	<a href="https://www.mitophen.org/">https://www.mitophen.org/</a>
<b>Sequenzdatenbanken</b>		
NCBI nucleotide database	Zusammenstellung von Sequenzen aus verschiedenen Datenbanken	<a href="https://www.ncbi.nlm.nih.gov/nucleotide/">https://www.ncbi.nlm.nih.gov/nucleotide/</a>
TTN-Transcript	Lokus Referenzgenom für TTN	<a href="https://www.cardiodb.org/titin/index.php">https://www.cardiodb.org/titin/index.php</a>

Aufgrund der Komplexität der Variantenbewertung wurden in den letzten Jahren zahlreiche Empfehlungen im Hinblick auf die Evidenzkriterien ausgesprochen. Auch in Zukunft werden weitere Spezifikationen zu erwarten sein. Über ein zugrundeliegendes Punktesystem zur

Bestimmung der fünf Klassen wird ebenfalls bereits diskutiert (28). Um die entsprechenden Kriterien und Empfehlungen auch zukünftig korrekt anzuwenden, sollte die Bewertung genetischer Veränderungen von humangenetischen Fachpersonal durchgeführt werden, um die Harmonisierung der Variantenbewertung vor dem Hintergrund der weiterhin zu erwartenden Komplexität zu gewährleisten. Eine Voraussetzung hierfür ist die Dokumentation des standardisierten Vorgehens bei der Variantenbewertung innerhalb einer Einrichtung. Auch eine kontinuierliche Schulung u.a. durch Konsensstrainings kann zur Harmonisierung beitragen.

### **3.6.2 Datenspeicherung**

Im Hinblick auf die Anforderungen zur Datenspeicherung in NGS-Prozessen sind verschiedene Aspekte zu berücksichtigen, die sich zum einen aus den gesetzlichen und normativen Vorgaben ergeben, zum anderen aber auch aus der in-house-Politik einer Einrichtung.

Wie bereits ausgeführt, gelten für die Genomdiagnostik dieselben gesetzlichen Vorgaben wie für die gesamte humangenetische Diagnostik. Insbesondere müssen auch bei der NGS-Datenspeicherung die allgemeine ärztliche Dokumentationspflicht für Patientenunterlagen von zehn Jahren nach Abschluss der Behandlung (Bürgerliches Gesetzbuch (BGB), § 630f BGB) sowie die Regelungen des GenDG (§12) beachtet werden. Letzteres schreibt die Aufbewahrung der Ergebnisse genetischer Untersuchungen und Analysen zehn Jahre in den Untersuchungsunterlagen der betroffenen Person vor. Die auch in diesem Paragraphen formulierte Vernichtungspflicht der Ergebnisse genetischer Untersuchungen und Analysen wird hier nicht weiter beschrieben werden, vielmehr sei auf S. Heidemann verwiesen (29).

Dabei ist zwischen den Ergebnissen genetischer Untersuchungen und Analysen und genetischen Daten zu unterscheiden. Während die Ergebnisse „eine technische Validierung oder Auswertung der genetischen Daten“ darstellen und für Ergebnisse einer übergeordneten genetischen Untersuchung eine abschließende Bewertung vorausgesetzt wird (3. GEKO-Tätigkeitsbericht,

<https://www.rki.de/DE/Content/Kommissionen/GendiagnostikKommission/>), sind die genetischen Daten diejenigen Daten, „die durch eine genetische Untersuchung [...] gewonnenen Daten über genetische Eigenschaften“ (GenDG§ 3 Nr. 11).

Während das GenDG für die Ergebnisse im Sinne der o.g. Ausführungen eine Aufbewahrungs- und auch Vernichtungspflicht vorsieht, werden diese für die genetischen Daten nicht definiert. Entsprechend hat S. Heidemann in ihrem Beitrag zur Vernichtungspflicht nach GenDG (29) darauf hingewiesen, dass in Bezug auf NGS-Daten im BCL-, FASTQ- oder BAM/CRAM-

Format diese nicht der Vernichtungspflicht unterliegen, die Vorgehensweise für VCF-Files dagegen zu diskutieren ist. Aus diesen Ausführungen ist abzuleiten, dass eine Dokumentationspflicht für NGS-Daten derzeit nicht eindeutig gefordert ist, da zumindest die Rohdaten zwar genetische Daten, aber nicht Ergebnisse im Sinne des GenDG darstellen.

Allgemein müssen die Dokumentationspflichten nach ISO15189 berücksichtigt werden, insbesondere im Hinblick auf Datensätze im Sinne von Aufzeichnungen (ISO15189: 4.13), die z.B. leicht zugänglich und gegen unbefugte Veränderungen geschützt sein müssen, aber auch im Falle von Änderungen mit Vermerk des Zeitpunktes und der ändernden Person versehen sein müssen.

Daher muss die Einrichtung, die NGS-Analysen anbietet, eine in-house-Politik zur NGS-Datenspeicherung formulieren, die neben den o.g. gesetzlichen Anforderungen weitere Aspekte berücksichtigt, wie fachlich-inhaltliche Aspekte, Anforderungen der Akkreditierungsnormen, und weitere Nutzung zur Beantwortung wissenschaftlicher Fragestellungen.

Die Entscheidung zum Umfang der Datenspeicherung sollte insbesondere berücksichtigen, inwieweit die verfügbaren und standardisierten Softwaremodule zur Datenprozessierung und -kompression zu einer fehlerhaften Auswertung führen können, und ob es zu einem für die Auswertung und Interpretation relevanten Verlust von Information kommen kann. Fällt diese Bewertung negativ aus, kann die Speicherung von VCF-Dateien ausreichend sein. Dabei ist aber zu berücksichtigen, ob nicht z.B. im Forschungskontext die Speicherung weiterer Dateiformate notwendig ist bzw. werden könnte. Dabei stellt die Speicherung von NGS-Daten als CRAM-Datei eine der platzeffizientesten Formen dar. Diese Dateien enthalten sowohl die Positionen der einzelnen Sequenzen im Genom, als auch fast alle Information der FASTQ-Dateien. Einzig Informationen hinsichtlich Rohdaten-Filterung, Kontrollnummer (die bei den modernen Sequenzierern nicht mehr genutzt wird) und der Indexsequenz, welche im Read Namen hinterlegt sind, gehen verloren. Außerdem wird die Information bezüglich Read 1 oder 2 anders als im FASTQ Format gespeichert. Sollten UMIs (unique molecular identifiers) zum Einsatz kommen, ist es notwendig diese in den Metadaten zu hinterlegen, da die im Read-Namen kodierte Information bei der Kompression verloren geht. Gegenüber FASTQ- und BAM-Dateien kann durch die Konvertierung zum CRAM-Format, eine Reduktion der Dateigröße von bis zu 60% erreicht werden, je nach gewünschtem Komprimierungsgrad mit oder ohne Datenverlust. Außerdem lassen sich in BAM- und CRAM- Dateien, im Gegensatz zu FASTQ-Dateien, zusätzliche Metadaten hinterlegen, sodass es möglich ist z.B. die Patientenennung direkt in der Datei einzubinden. Obwohl sich CRAM-Dateien zunehmend zum bevorzugten Datenspeicherungsformat entwickeln, müssen die gewählten

Softwarelösungen für die Sekundär- und Tertiär-Analyse effizient mit dem Format umgehen können. Daher konvertieren viele Labors CRAM- zunächst in BAM-Dateien und verwenden CRAM-Dateien nur für Archivierungszwecke (30).

Vor diesem Hintergrund bleibt es dem Labor überlassen, eine Regelung zur Speicherung von NGS-Datensätzen zu formulieren. So kann eine Speicherung der Variantentabellen/VCF-Dateien ausreichend sein, da auf Basis der Variantentabellen die anschließenden Auswertungs- und Interpretationsschritte nachvollzogen werden können, die zur Generierung der Ergebnisse im Sinne des GenDG beitragen. Dabei ist aber zu berücksichtigen, dass die ausschließliche Speicherung von VCF-Dateien eine Überprüfung von Varianten zu einem späteren Zeitpunkt, z.B. in Hinblick auf das Vorhandensein von Artefakten wie Alignmentfehlern, unmöglich macht. Eine visuelle Inspektion von Varianten z.B. mittels IGV Viewer ist somit nicht mehr möglich. Dies kann eine Einschränkung bei der periodischen Re-Analyse von NGS-Daten darstellen.

Sollte eine Analyse von Rohdaten erneut erforderlich sein, ist prinzipiell eine erneute NGS-Analyse (Nasslabor, Datenprozessierung), ggf. mit erneuter Probengewinnung, geeignet, um auch methodische Weiterentwicklungen der NGS-Diagnostik zu nutzen. Allerdings ist dies nicht immer möglich, so dass jede Einrichtung prüfen sollte, auch die Sequenzdaten selbst in Form von FASTQ, BAM oder CRAM-Dateien, zu archivieren.

### **3.7 Befundberichte**

Der Aufbau und Inhalt von Befundberichten richtet sich nach den Anforderungen der ISO15189 und der S2k-Leitlinie Humangenetische Diagnostik (20). Neben den Patientenidentifikationsmerkmalen und der Indikation sollte das für den Einsender relevante Ergebnis in Hinblick auf Diagnose, Prädiktion, Therapie oder genetische Beratung prägnant möglichst auf der ersten Seite beschrieben werden (EuroGen recom. 38 (5)). Wie bei anderen Befundberichten zu genetischen Analysen muss auch die Methode eindeutig beschrieben sein und alle wesentlichen Auswerteparameter enthalten, insbesondere Angaben zum Testverfahren, den mit der Methode nachweisbaren Variantentypen, die Abdeckung und Lesetiefe der Zielregion (region of interest), d.h. des Exoms, Genoms, und/oder der Gene, über die explizit berichtet wird, sowie zur analytischen Sensitivität und Präzision (EuroGen recom. 37 (5)). Angaben zum Referenzgenom, Referenz-Transkript-Version und OMIM-Referenz gehören ebenso in den Befund (EuroGen recom. 39, 40 (5)) wie Hinweise auf die Variantenbewertung (z.B. ACMG). Als Referenztranskripte sollten die von der MANE-Initiative vorgeschlagenen

Versionen genutzt werden. Inwieweit zur weiteren Erläuterung der klinischen Relevanz (wahrscheinlich) pathogener Varianten Datenbank- und Literaturhinweise im Befund erwähnt werden, ist durch das Labor zu prüfen. Während pathogene und wahrscheinlich pathogene Varianten nach ACMG-Einteilung, die in Zusammenhang mit dem Untersuchungsauftrag stehen, berichtet werden, sollten VUS nur in Genen ausgewertet und berichtet werden, die mit einem bekannten Phänotyp assoziiert sind (EuroGen recom. 42 (5), (2, 6)).

### **3.8 Qualitätssichernde Maßnahmen**

Das Laboratorium muss Untersuchungsverfahren auswählen und validieren bzw. verifizieren, die für die Abklärung der medizinischen Fragestellung geeignet sind. Zur Validierung NGS-basierender Untersuchungsverfahren stehen Leitlinien und Hinweise zur Verfügung (2, 5, 6). Für die GS wurden ergänzende spezielle Empfehlungen publiziert (21). Die Validierungen der Plattform und der Pipeline sollten Referenzmaterialien wie Genome-in-a-bottle Proben (GiaB, Tab. 4) berücksichtigen. Die Verwendung von Genpanels für die Abklärung gezielter Fragestellungen impliziert eine Validierung auf Genebene. Dabei sollte die Abdeckung der bekannten krankheitsverursachenden Varianten gewährleistet und mögliche Limitationen der Methodik (z. Bsp. Pseudogene,) berücksichtigt werden. Als Referenzsequenzen sollten die mittlerweile von Datenbanken harmonisierte Transkripte verwendet werden (MANE (3)).

Für die Probenverfolgung sollte eine Genotypisierung mit einer orthogonalen Methode etabliert sein, über die SNVs abgeglichen werden (EuroGen recom. 29 (5)), wenn die Probenverfolgung nicht anderweitig gesichert ist. Die Qualitätskontrolle der Analysen sollte mindestens die Überwachung der vom Labor festgelegten Coverage umfassen.

Bei pränatalen Analysen gehört der Ausschluss einer maternalen Kontamination ebenfalls zu den obligaten qualitätssichernden Maßnahmen.

Zur Sicherstellung der Qualität der Untersuchungen gehört auch die Teilnahme an externen qualitätssichernden Maßnahmen (Ringversuchen), durch die das Erreichen der vorgesehenen Qualität nachgewiesen wird. Die Anforderungen werden in der derzeit im Entwurf befindlichen neuen RiliBÄK-Version geregelt, die ISO15189 macht allgemeine Vorgaben zur Qualitätssicherung. Das Laboratorium muss Ringversuchsprogramme auswählen, die die klinisch relevanten Fragestellungen abdecken und jeweils das gesamte Untersuchungsverfahren überprüfen. Hierzu zählen der Nasslaborbereich, die Datenprozessierung, die Variantenklassifikation und -interpretation als auch die Befundung. Entsprechende Ringversuche werden z. Bsp. beim BVDH, EMQN, UK NEQAS und GenQA angeboten.



## Links:

ACGS: <https://www.acgs.uk.com/media/11631/uk-practice-guidelines-for-variant-classification-v4-01-2020.pdf>

ClinGen: <https://clinicalgenome.org/working-groups/sequence-variant-interpretation/>

## RiliBÄK:

[https://www.bundesaerztekammer.de/fileadmin/user\\_upload/BAEK/Themen/Qualitaetsicherung/\\_Bek\\_BAEK\\_RiLi\\_QS\\_laboratoriumsmedizinischer\\_Untersuchungen.pdf](https://www.bundesaerztekammer.de/fileadmin/user_upload/BAEK/Themen/Qualitaetsicherung/_Bek_BAEK_RiLi_QS_laboratoriumsmedizinischer_Untersuchungen.pdf)

## Literatur:

1. Palmquist R, Jenkins SM, Bentley D, Miller C, Mao R, Meibos B, et al. Evaluating use of changing technologies for rapid next-generation sequencing in pediatrics. *Pediatr Res.* 2022.
2. Matthijs G, Souche E, Alders M, Corveleyn A, Eck S, Feenstra I, et al. Guidelines for diagnostic next-generation sequencing. *Eur J Hum Genet.* 2016;24(1):2-5.
3. Morales J, Pujar S, Loveland JE, Astashyn A, Bennett R, Berry A, et al. A joint NCBI and EMBL-EBI transcript set for clinical genomics and research. *Nature.* 2022;604(7905):310-5.
4. Costain G, Walker S, Marano M, Veenma D, Snell M, Curtis M, et al. Genome Sequencing as a Diagnostic Test in Children With Unexplained Medical Complexity. *JAMA Netw Open.* 2020;3(9):e2018109.
5. Souche E, Beltran S, Brosens E, Belmont JW, Fossum M, Riess O, et al. Recommendations for whole genome sequencing in diagnostics for rare diseases. *Eur J Hum Genet.* 2022;30(9):1017-21.
6. S1 Leitlinie: Molekulargenetische Diagnostik mit Hochdurchsatz-Verfahren der Keimbahn, beispielsweise mit Next-Generation Sequencing. *Medizinische Genetik.* 2018;30(2):278-92.
7. Marshall CR, Chowdhury S, Taft RJ, Lebo MS, Buchan JG, Harrison SM, et al. Best practices for the analytical validation of clinical whole-genome sequencing intended for the diagnosis of germline disease. *NPJ Genom Med.* 2020;5:47.
8. Krusche P, Trigg L, Boutros PC, Mason CE, De La Vega FM, Moore BL, et al. Best practices for benchmarking germline small-variant calls in human genomes. *Nat Biotechnol.* 2019;37(5):555-60.
9. Rüping U, Stock N, Schmidtke J. Ärztliche Aufklärungspflichten über die Möglichkeit von Nebenbefunden in der genomischen Medizin. *Medizinrecht.* 2022;40(8):663-7.
10. Richtlinie der Gendiagnostik-Kommission (GEKO) für die Anforderungen an die Inhalte der Aufklärung bei genetischen Untersuchungen zu medizinischen Zwecken gemäß § 23 Abs. 2 Nr. 3 GenDG. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz.* 2022;65(9):963-8.
11. Maxwell KN, Hart SN, Vijai J, Schrader KA, Slavin TP, Thomas T, et al. Evaluation of ACMG-Guideline-Based Variant Classification of Cancer Susceptibility and Non-Cancer-Associated Genes in Families Affected by Breast Cancer. *Am J Hum Genet.* 2016;98(5):801-17.

12. Kalia SS, Adelman K, Bale SJ, Chung WK, Eng C, Evans JP, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. *Genet Med*. 2017;19(2):249-55.
13. Webber EM, Hunter JE, Biesecker LG, Buchanan AH, Clarke EV, Currey E, et al. Evidence-based assessments of clinical actionability in the context of secondary findings: Updates from ClinGen's Actionability Working Group. *Hum Mutat*. 2018;39(11):1677-85.
14. Kraft F, Kurth I. Long-read sequencing in human genetics. *medizinische genetik*. 2019;31(2):198-204.
15. Herzog TJ, Vergote I, Gomella LG, Milenkova T, French T, Tonikian R, et al. Testing for homologous recombination repair or homologous recombination deficiency for poly (ADP-ribose) polymerase inhibitors: A current perspective. *Eur J Cancer*. 2023;179:136-46.
16. Abicht A, Scharf F, Kleinle S, Schon U, Holinski-Feder E, Horvath R, et al. Mitochondrial and nuclear disease panel (Mito-aND-Panel): Combined sequencing of mitochondrial and nuclear DNA by a cost-effective and sensitive NGS-based method. *Mol Genet Genomic Med*. 2018;6(6):1188-98.
17. McCormick EM, Lott MT, Dulik MC, Shen L, Attimonelli M, Vitale O, et al. Specifications of the ACMG/AMP standards and guidelines for mitochondrial DNA variant interpretation. *Hum Mutat*. 2020;41(12):2028-57.
18. Eggermann K, Meyer R, Begemann M, Dey D, Bultmann E, Kurth I, et al. Clonal Elimination of the Pathogenic Allele as Diagnostic Pitfall in SAMD9L-Associated Neuropathy. *Genes (Basel)*. 2022;13(12).
19. Deans ZC, Allen S, Jenkins L, Khawaja F, Gutowska-Ding W, Patton SJ, et al. Ensuring high standards for the delivery of NIPT world-wide: Development of an international external quality assessment scheme. *Prenat Diagn*. 2019;39(5):379-87.
20. S2k-Leitlinie Humangenetische Diagnostik und Genetische Beratung. *medizinische genetik*. 2018;30(4):469-522.
21. Wagner J, Olson ND, Harris L, McDaniel J, Cheng H, Functamman A, et al. Towards a Comprehensive Variation Benchmark for Challenging Medically-Relevant Autosomal Genes. *bioRxiv*. 2021:2021.06.07.444885.
22. Aganezov S, Yan SM, Soto DC, Kirsche M, Zarate S, Avdeyev P, et al. A complete reference genome improves analysis of human genetic variation. *Science*. 2022;376(6588):eabl3533.
23. Rakocevic G, Semenyuk V, Lee WP, Spencer J, Browning J, Johnson IJ, et al. Fast and accurate genomic analyses using genome graphs. *Nat Genet*. 2019;51(2):354-62.
24. Koboldt DC. Best practices for variant calling in clinical sequencing. *Genome Medicine*. 2020;12(1):91.
25. Roy S, Coldren C, Karunamurthy A, Kip NS, Klee EW, Lincoln SE, et al. Standards and Guidelines for Validating Next-Generation Sequencing Bioinformatics Pipelines: A Joint Recommendation of the Association for Molecular Pathology and the College of American Pathologists. *J Mol Diagn*. 2018;20(1):4-27.
26. Hardwick SA, Deveson IW, Mercer TR. Reference standards for next-generation sequencing. *Nat Rev Genet*. 2017;18(8):473-84.
27. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405-24.
28. Tavtigian SV, Harrison SM, Boucher KM, Biesecker LG. Fitting a naturally scaled point system to the ACMG/AMP variant classification guidelines. *Hum Mutat*. 2020;41(10):1734-7.

29. Heidemann S. 10 Jahre Gendiagnostikgesetz – wie kann die Vernichtungspflicht für Ergebnisse genetischer Analysen und Untersuchungen praktisch umgesetzt werden? *Medizinische Genetik*. 2020;32(1):75-8.
30. Bonfield JK. CRAM 3.1: advances in the CRAM file format. *Bioinformatics*. 2022;38(6):1497-503.



**Florian Kraft**  
 Institut für Humangenetik und Genommedizin  
 Medizinische Fakultät der RWTH Aachen  
 Aachen  
 Deutschland  
**e-mail:** [fkraft@ukaachen.de](mailto:fkraft@ukaachen.de)  
**https://orcid.org/0000-0002-5324-9155**



**Peter Bauer**  
 CENTOGENE N.V.  
 Rostock  
 Deutschland  
**e-mail:** [peter.bauer@centogene.com](mailto:peter.bauer@centogene.com)



**Anna Benet-Pagès**  
 Institut für Neurogenomik  
 Helmholtz Zentrum München  
 Neuherberg  
 Deutschland  
 Medizinisch Genetisches Zentrum  
 München, Deutschland  
**e-mail:**  
**[Anna.Benet-Pages@mgz-muenchen.de](mailto:Anna.Benet-Pages@mgz-muenchen.de)**



**Matthias Begemann**  
 Institut für Humangenetik und Genommedizin  
 Medizinische Fakultät der RWTH Aachen  
 Aachen  
 Deutschland  
**e-mail:** [mbegemann@ukaachen.de](mailto:mbegemann@ukaachen.de)  
**https://orcid.org/0000-0002-4659-8437**



**Daniel Berner**  
 MVZ genetikum GmbH  
 Neu-Ulm  
 Deutschland  
**e-mail:** [berner@genetikum.de](mailto:berner@genetikum.de)



**Marc Sturm**  
 Institut für Medizinische Genetik und Ange-  
 wandte Genomik  
 Universität Tübingen  
 Tübingen  
 Deutschland  
**e-mail:** [Marc.Sturm@med.uni-tuebingen.de](mailto:Marc.Sturm@med.uni-tuebingen.de)



**Anna Teubert**  
 amedes genetics  
 Hannover  
 Deutschland  
**e-mail:** [anna.teubert@amedes-group.com](mailto:anna.teubert@amedes-group.com)



**Stephanie Kleinle**  
 Medizinisch Genetisches Zentrum  
 München  
 Deutschland  
**e-mail:**  
**[stephanie.kleinle@mgz-muenchen.de](mailto:stephanie.kleinle@mgz-muenchen.de)**



**Sebastian Eck**  
 Bionano  
 San Diego  
 CA 92121  
 USA  
**e-mail:** [sebastianeck@web.de](mailto:sebastianeck@web.de)



**Tobias B. Haack**  
 Institut für Medizinische Genetik und Ange-  
 wandte Genomik  
 Universität Tübingen  
 Tübingen  
 Deutschland  
**e-mail:**  
**[tobias.haack@med.uni-tuebingen.de](mailto:tobias.haack@med.uni-tuebingen.de)**



**Norbert Arnold**  
 Zentrum für familiären Brust- und Eierstock-  
 krebs  
 Klinik für Gynäkologie und Geburtshilfe  
 Universitätsklinikum Schleswig-Holstein  
 Kiel  
 Deutschland  
**e-mail:** [norbert.arnold@uksh.de](mailto:norbert.arnold@uksh.de)



**Prof. Dr. rer. nat. Thomas Eggermann**  
 Institut für Humangenetik und Genommedizin  
 Medizinische Fakultät der RWTH Aachen  
 Pauwelsstr. 30, D-52074 Aachen  
 Germany  
 Tel.: 0241 8037285  
 Fax:0241 8082580  
**e-mail:** [teggermann@ukaachen.de](mailto:teggermann@ukaachen.de)