

## Doporučení Eurogentestu pro diagnostické NGS

Spiros Tavandzis a kolektiv

Text byl převzat a upraven z originálu „Guidelines for diagnostic next generation sequencing, Eurogentest 2014“. Překlad byl především využit pro validaci NGS platformy, některé části textu byly vynechány. Prosíme proto čtenáře, aby text považovali za stručný přehled doporučení. Budeme rádi za Vaše komentáře ([spiros.tavandzis@lag.agel.cz](mailto:spiros.tavandzis@lag.agel.cz)).

### Slovník základních pojmů:

**diagnostická výtěžnost** (diagnostic yield)- pravděpodobnost, že bude nalezena varianta asociována s onemocněním a že bude moci být stanovena diagnóza. Její hodnota je určována na základě souboru vyšetřovaných pacientů. Stanovuje výkonnost NGS především z klinického hlediska.

**klinický cíl - oblast zájmu (region of interest - ROI)** – slouží k vymezení vykazovaného rozsahu a návrhu dg. testu

**vykazovaný rozsah** – oblast klinických cílů (ROIs), kterou lze spolehlivě sledovat např. bez drop outů apod.

**technický cíl** (technical target) – podíl duplikátů, množství dat, pokrytí, počet clusterů atd.

**pokrytí (coverage)** – počet identifikací konkrétní báze na konkrétní pozici

**calls** - identifikace báze na konkrétní pozici

**Informativní pokrytí (informative coverage - IC)** - podíl cíle, u něhož může být spolehlivě stanoven genotyp. Liší se od běžného (hrubého) pokrytí.

**pipeline** - bioinformatický analytický postup

**sekundární nález** - nález v genu, který **není** zapojen v etiologii testovaného onemocnění, ale je součástí panelu

**nevyžádaný nález** - nález zachycený v genu zapojeném v etiologii testovaného onemocnění, nejedná se však o klíčový gen

**klíčové geny (core genes)** – **klinicky relevantní geny, geny prokazatelně** asociované s daným onemocněním

### Prohlášení:

**1.01:** NGS by nemělo být zavedeno do klinické praxe bez přípustné validace provedené v souladu s nově vznikajícími doporučeními.

**1.02:** Laboratoř musí jasně deklarovat, že nabízený test diagnózu vylučuje nebo potvrzuje.

**2.01:** Cíl a přínos testu by měly být projednány na začátku validace. Shrnutí cíle a přínosu by mělo být součástí validační zprávy.

**2.02:** Pokud laboratoř zvažuje zavedení NGS do diagnostiky, měla by nejprve posoudit jeho „diagnostickou výtěžnost“

**2.03:** Pro diagnostické účely by měly být v analýze zahrnuty pouze geny se známým vztahem (t.j. publikovaným a potvrzeným) mezi aberantním genotypem a patologií daného onemocnění.

**2.04:** Klinickými a laboratorními odborníky by měl být vytvořen „seznam klíčových genů daného onemocnění“ z několika důvodů: pro účely porovnání, zabránění nezodpovědnému/nespolehlivému testování, prospěšnost pro pacienta.

**2.05:** Porovnání nabídek diagnostického testování mezi laboratořemi by měl umožnit jednoduchý hodnotící systém na základě pokrytí a diagnostického přínosu (výtěžku).

**3.01:** Laboratoř musí pro každý NGS test uvádět: onemocnění, na které/která je test zaměřen; názvy testovaných genů; jejich vykazovaný rozsah; analytickou citlivost a specifitu testu; a pokud je to možné i onemocnění, která nejsou ke klinickému fenotypu relevantní, která by ale mohla být způsobena mutacemi v testovaných genech.

**3.02:** Aby se snížila možnost sekundárních nálezů, měly by se diagnostické laboratoře ve svých analytických postupech soustředit na to, jaké geny jsou součástí panelu, a na jejich odpovídající validaci.

**3.03:** Laboratoře by měly poskytovat informace o možnosti nevyžádaných nálezů.

**3.04:** Pokud se klinické centrum nebo laboratoř rozhodne pacientům podávat informaci o přenašečství u nesouvisejícího onemocnění a o sekundárních nálezech, měly by zavést tzv. opt-in/opt-out protokoly s veškerými náležitostmi.

**3.05:** Pacientovi by měla být jasná pravidla o poskytování nevyžádaných a sekundárních nálezů dané laboratoře.

**3.06:** Doporučuje se poskytnout pacientovi informace písemně formou letáku, nebo online na webových stránkách.

**4.01:** Všechna měření kvality používaná při NGS diagnostických postupech by měla být přesně popsána.

**4.02:** Diagnostické laboratoře musí zavést strukturovanou databázi relevantních měření kvality v rámci (i) všech typů platform (ii) všech typů testů (iii) a všech zpracovaných vzorků.

**4.03:** Součástí nastavení celé metody musí být i jednoznačná identifikace vzorků a možnost sledování (z hlediska nezaměnitelnosti vzorků, v originále zmiňují využití čárových kódů) v celém průběhu jejich zpracování. Tyto aspekty musí být zahrnuty ve validaci platformy.

**4.04:** Přesnost a preciznost (accuracy and precision) by měly být součástí obecné validace platformy. Validace nemusí být pro jednotlivé metody nebo testy opakována.

**4.05:** Bioinformatická analýza musí být upravena pro konkrétní použitou technickou platformu.

**4.06:** Analytická citlivost a specifita musí být v průběhu validace analýzy stanoveny samostatně pro každý typ varianty.

**4.07:** Diagnostická laboratoř musí validovat všechny části bioinformatického analytického postupu pomocí standardizovaných datových souborů kdykoli, kdy jsou implementovány relevantní změny – nové verze, aktualizace (platí pro veřejné online nástroje i komerční softwarové balíky).

**4.08:** Diagnostická laboratoř musí používat strukturovanou databázi všech relevantních variant s aktuálními anotacemi.

**4.09:** Diagnostická laboratoř musí zařídit dlouhodobé uložení všech relevantních dat.

**4.10:** Během vývoje testu musí být definován „vykazovaný rozsah“, což je část „regions of interest“ (ROI), pro kterou je spolehlivě generován určitý počet „volání“ („calls“, VCF). Vykazovaný rozsah by měl být k dispozici lékařům (např. ve zprávě, nebo online).

**4.11:** Požadavky na „vykazovaný rozsah“ závisí na cíli testu.

**4.12:** Kdykoliv dojde k důležitým změnám v metodě, musí být zkontrolovány parametry kvality a reanalyzovány vzorky. Laboratoř by měla předem stanovit jaké vzorky a kolik jich bude reanalyzováno pokud k aktualizaci metody dojde.

**5.01:** Zpráva z NGS testování by měla na jedné stránce obsahovat identifikaci pacienta, diagnózu, stručný popis testu, shrnutí výsledků a hlavní nález.

**5.02:** Dříve, než laboratoř poskytne analýzu tohoto typu, měla by nastavit a zdokumentovat pravidla, podle kterých bude reportovat genomické varianty. Pravidla by měla být v souladu s mezinárodními doporučeními.

**5.03:** Údaje o variantách nejasného významu (VUS, UV) musí být shromažďovány z důvodů případné definitivní klasifikace jejich významu.

**5.04:** Laboratoře by měly mít před spuštěním testu jasně definovaný postup/protokol pro případ řešení nevyžádaných a sekundárních nálezů.

**5.05:** Neočekává se, že by laboratoř systematicky re-analyzovala stará data a reportovala nové nálezy, a to ani v případě, že dojde ke změnám v panelech klíčových genů.

**5.06:** Laboratoř musí vytvořit vlastní databáze variant pro onemocnění, jejichž testování nabízí, aby byla schopna varianty vyhodnocovat.

**6.01:** Diagnostický test je jakýkoliv test prováděný za účelem odpovědi na otázku související se zdravotním stavem pacienta.

**6.02:** Výzkumný test je založen na hypotéze a výsledek může mít pro pacienta zapojeného do projektu omezený klinický význam.

**6.03:** Výsledky diagnostického testu mohou vést k hypotéze.

**6.04:** Diagnostické testy, jejichž primárním cílem je hledání diagnózy konkrétního pacienta, by měly být prováděny akreditovanými laboratořemi.

**6.05:** Před předložením výsledků výzkumu indikujícímu lékaři a/nebo pacientovi (klientovi), musí být tyto výsledky potvrzeny v akreditované laboratoři.

**6.06:** Frekvence všech variant, které jsou detekovány sekvenováním vzorků zdravých jedinců v diagnostickém a/nebo výzkumném režimu, by měly být sdíleny.

**6.07:** Všechny varianty uváděné ve výsledkových zprávách by měly být sdíleny v regionálních, národních i mezinárodních databázích.

## Kapitola 1: Obecný úvod

### 1.1 Úvod

**Prohlášení 1.01: NGS by nemělo být zavedeno do klinické praxe bez odpovídající validace provedené v souladu s nově vznikajícími doporučeními.**

Ať už je cílem diagnostického testu vyloučit, nebo potvrdit diagnózu, je nutno tuto skutečnost definovat předem, protože se jedná o významný rozdíl. Rozdíl je především v kompletnosti testu a podmiňuje nejen různá nastavení, ale především jiný pohled na diagnostiku.

Pokud je prováděno NGS subdodavatelem (jako servisní služba), je esenciální dosáhnout stejných podmínek kvality, jako v případě in-house sekvenování. Doporučujeme, aby byl poskytovatel takové služby akreditovaný oprávněným akreditačním institutem, a aby byla servisní smlouva jasně definovaná a navržena tak, aby garantovala provedení testování v souladu s akreditačními standardy pro diagnostické testování (ISO 15189).

Celý tento text Doporučení se v podstatě zabývá NGS testováním v souvislosti se vzácnými a většinou monogenními nemocemi. Aplikovatelné je i na somatické testování v souvislosti s nádorovým onemocněním. Testování somatických variant ale vyžaduje další parametry kvality, jako je práh detekce, což se obecně u varianty testované v germinální linii neřeší. Tento dokument tuto problematiku nezmiňuje.

Doporučení se zabývá hlavně analýzou pomocí panelů genů a to specifickou capture assay, nebo exomovým sekvenováním. Důvody tohoto přístupu byly nedávno publikovány jinde a nebudou zde opakovány (Rehm 2013). Pro celogenomové sekvenování (whole genome sequencing, WGS) jsou tato doporučení v zásadě také platná, je ale nutné je odpovídajícím způsobem rozšířit.

Využití NGS pro určování rizikových faktorů u multifaktoriálních onemocnění není v současnosti klinicky akceptovaná praxe. Proto ani neuvádíme specifika pro nabízení takového druh testování.

**Prohlášení 1.02: Laboratoř musí jasně deklarovat, že nabízený test diagnózu vylučuje nebo potvrzuje.**

Takové rozlišení je důležité a zaručuje odlišné nastavení a rozdílný pohled na diagnostiku. Podobně, pokud laboratoře nabízí somatické testování pomocí NGS, měly by být jasně definovány limity této metody.

### 1.2 Vývoj doporučení pro diagnostické využití

#### 1.2.1 Rozsah

Platforma masivního paralelního sekvenování je používána pro různé aplikace. V rámci diagnostiky rozlišujeme následující NGS assaye:

- mutation scanning (využívá se pro konkrétní gen nebo malou skupinu genů).
- mutační screening pomocí „target capture“ nebo amplikonovým sekvenováním známých genů.

- exomové sekvenování by mělo být rozděleno na dvě odlišné aplikace. První je cílená analýza známých genů, kdy jsou pokyny stejné jako v případě cíleného mutačního screeningu s výjimkou v případě nevyžádaných výsledků – v informovaném souhlasu to musí být popsáno rozsáhleji. Druhá aplikace je identifikace nových genetických poruch, což podle nás z větší části zůstává v oblasti výzkumu hlavně v případě, kdy geny, ve kterých byla varianta nalezena, nebyly s konkrétním onemocněním dříve spojovány; t.j. je obtížné takovou podrobnou analýzu v diagnostice nabízet. Výjimkou z takového postoje je exomové sekvenování pacient + rodiče z důvodu identifikace *de novo* vad.

- „mendeliomy“ (kombinace cíleného a exomového testování)

- celogenomové sekvenování (Whole genome sequencing – WGS)

Každá platforma má svá technická omezení (Buermans a den Dunnen, 2014) i konceptuální limity – např. trinukleotidové repetice nemohou být sekvenováním a mapováním krátkých úseků detekovány. Tato doporučení ale nepopisují jednotlivé platformy a jejich specifické technické parametry. Laboratoř musí tato omezení a limity znát a v rutinním diagnostickém testování zohlednit. O limitech daného diagnostického testu je důležité informovat klienty (klinické pracovníky).

### 1.2.2 Metody

- nepřeloženo, obsahuje výčet dalších starších doporučení

### 1.2.3 Limity

Tato doporučení nehodnotí výhody a nevýhody diagnostiky s použitím „target capture assay“ vs. exomové sekvenování. Je odpovědností diagnostické laboratoře toto vyhodnotit a obhájit v detailní validační zprávě, před zavedením metody. Tato doporučení jsou obecná – není možné popsat všechny oblasti aplikace do všech možných detailů.

Doporučení jsou výsledkem diskuze relativně malé skupiny expertů, která nebyla tvořena zástupci všech subjektů. Nicméně ambicí těchto doporučení je, aby byla převzata národními akreditačními institucemi k doplnění normy ISO 15189.

### 1.2.4 Přínos EuroGentestu

- nepřeloženo

## 1.3 Hlavní body dokumentu

1. Doporučení zdůrazňují povinnost laboratoře před zahájením nabídky NGS diagnostického testu, definovat jeho diagnostickou výtěžnost
2. Nejdůležitějším novým prvkem tohoto dokumentu je návrh na hodnocení různých NGS testů jako typy A, B nebo C v závislosti na jejich kvalitě a komplexnosti
3. Autoři navrhují standardizaci parametrů kvality s využitím „vykazovaného rozsahu“
4. Laboratoř musí přijmout opatření, která umožní vypořádat se s dalšími zvláštnostmi, které jsou dané povahou NGS testování, jako jsou sekundární a nevyžádané nálezy nebo zjištění přenašečství u recesivních nebo X-vázaných onemocnění.

Je nad rámec úkolů vedoucích laboratoří definovat, jaký by měl být obsah a využití informovaného souhlasu. Proto je v tomto dokumentu uvedeno, že není zodpovědností laboratoře informovaný souhlas poskytovat ani uchovávat.

5. Doporučuje se striktně rozlišovat mezi diagnostikou a výzkumem, přesto, že se tyto rozdíly díky NGS technologii stírají.

## Kapitola 2: Diagnostické/klinické využití

### 2.1 Úvod

Výhody zavedení NGS do rutinní diagnostiky jsou v testování mnoha genů najednou v relativně krátkém čase za relativně nízkou cenu, čímž lze pokrýt více diagnóz najednou, a to exomovým nebo genomovým sekvenováním, nebo cílenou analýzou vybrané skupiny genů.

V textu je složitě popisováno, že v případě cílené analýzy má panel genů buď obsahovat geny asociované s daným onemocněním (klíčové geny), nebo mají být geny asociované s konkrétním onemocněním vyfiltrovány v průběhu bioinformatické analýzy.

Dále z důvodu porovnatelnosti nabídky jednotlivých laboratoří navrhují autoři pro diagnostické NGS testy zavedení tzv. „rating scheme“ pro jehož vytvoření je nutná národní i mezinárodní spolupráce.

### 2.2 Hlediska a příklady

#### 2.2.1 Limity NGS a diagnostická výtěžnost

Limity NGS závisí na platformě a obohacovací metodě - enrichmentu (pokud je použita). Enrichment založený na PCR je citlivý na alelické dropouty způsobené SNPs v místě nasedání primerů, ale je levnější a méně pracný a snadno aplikovatelný na malý počet pacientů než metody založené na capture. Capture assay nemají takový problém s alelickým dropoutem, ale neporadí si dostatečně s oblastmi bohatými na GC. WGS je méně zatíženo alelickým dropoutem, ale vyžaduje vysokou sekvenační kapacitu náhradou za nižší coverage, atd.

**Prohlášení 2.01: Cíl a přínos testu by měly být projednány na začátku validace. Shrnutí cíle a přínosu by mělo být součástí validační zprávy.**

**Prohlášení 2.02: Pokud laboratoř zvažuje zavedení NGS do diagnostiky, měla by nejprve posoudit jeho „diagnostickou výtěžnost“**

"Diagnostická výtěžnost" je definována jako pravděpodobnost, že bude nalezena varianta asociovaná s onemocněním a že bude moci být stanovena diagnóza. Její hodnota je určována na základě souboru vyšetřovaných pacientů (Weiss, Van der Zwaag et al., 2012). Stanovuje výkonnost NGS z klinického hlediska. „Diagnostickou výtěžnost“ není jednoduché určit, protože definovat soubor pacientů pro danou diagnózu nebo klinický požadavek může být komplikované. Může být ale využita literatura a porovnání s existujícími technikami (Neveling et al., 2013).

Na základě „diagnostické výtěžnosti“ by měl být vytvořen „seznam klíčových genů“ a „diagnostická cesta/směrování“. „Diagnostická výtěžnost“ není parametr kvality laboratoře, těmi jsou citlivost a specifita.

#### 2.2.2 Seznam klíčových genů

Vytvářené panely pro konkrétní onemocnění mají obsahovat geny, které jsou prokazatelně zodpovědné za vznik onemocnění. Tyto geny lze označit jako klíčové. Dobrovolně lze do panelu zařadit také geny, které s daným onemocněním souvisejí „méně“.

**Prohlášení 2.03: Pro diagnostické účely by měly být v analýze zahrnuty pouze geny se známým vztahem (tj. publikovaným a potvrzeným) mezi aberantním genotypem a patologií daného onemocnění.**

Další věc je nastavení standardu pro coverage a citlivost. Mělo by být určeno, pro které geny platí stejná analytická citlivost jako v případě Sangerova sekvenování. Při přechodu ze Sangerova sekvenování na NGS by neměla být analytická citlivost vyšetření pro daný gen snížena. Přidání dalších genů do panelu bude samozřejmě zvyšovat „diagnostickou výtěžnost“, ale nemělo by to být na úkor ztrát mutací, které by byly dříve detekovány.

Je nezbytné, aby byly úseky s nízkým pokrytím dovyšetřeny jinou metodou – např. Sangerovým sekvenováním. Pokud ale v úsecích, které mají nízké pokrytí (a byly dosekvenovány Sangerem), nebyla u stovek až tisíců případů identifikována žádná mutace, nemá smysl takový úsek dosekvenovávat. Laboratoř ale musí tuto skutečnost doložit na základě vlastních zkušeností nebo na základě publikovaných dat.

Vzrůstající detekční poměr (záchytnost) je při definování seznamu klíčových genů a vypořádání se s nepokrytými úseky hlavním faktorem.

V souhrnu lze konstatovat, že představy o seznamu klíčových genů jsou následující:

- seznam musí podstatnou měrou přispět ke kvalitě života pacienta a proto musí být geny vybrány pečlivě
- byl by přijatelný dvoustupňový systém, kde jsou některé geny zkoumány podrobněji (jinými slovy, s více kompletním pokrytím) než jiné
- seznam nesmí působit na účinnost poskytované služby tj. příliš horlivé testování není užitečné
- použití panelů klíčových genů musí vést k lepší diagnóze skupiny onemocnění, pokud ne, pak nemá klinickou prospěšnost

**Prohlášení 2.04: Klinickými a laboratorními odborníky by měl být vytvořen „seznam klíčových genů daného onemocnění“ z několika důvodů: pro účely porovnání, zabránění nezodpovědnému/nespolehlivému testování, prospěšnost pro pacienta.**

Nabízené testy by měly být klinicky relevantní.

### 2.2.3 NGS vs. jiné techniky: „diagnostická cesta/směrování“

„Diagnostická cesta/směrování“ vyjadřuje celý diagnostický přístup laboratoře k danému genu. Zda-li jej vyšetřuje pomocí NGS, dosekvenovává oblasti s nízkou coverage, dovyšetřuje gen metodou MLPA, CGH array, WGS atd. Diagnostický test by měl být tvořen NGS s následným dovyšetřením.

Laboratoře budou aplikovat různé (technické a diagnostické) nastavení NGS testů, a to i bez ohledu na pokyny. Existuje totiž stále příliš mnoho proměnných, které nemohou být přikazovány prostřednictvím normativních pokynů. Proto navrhujeme **jednoduchý systém hodnocení (laboratoře typu A, B, C)** pro NGS diagnostiku, který bude zárukou spravedlivého hodnocení a snadného porovnání nabídek různých laboratoří.

1. **Test typu A:** Laboratoř zaručuje >99% spolehlivost identifikace bází (calls) referenční sekvence nebo variant v kódující oblasti a přilehlých intronových sekvencích, kompletně dosekvenovává výpadky Sangerovým nebo srovnatelným sekvenováním a v závislosti na použité platformě provádí další analýzy, například dosekvenování homopolymerních oblastí. Toto je nejvyšší úroveň přesnosti, kterou může laboratoř v současné době u NGS nabídnout.
2. **Test typu B:** Laboratoř přesně udává, které úseky jsou sekvenovány s >99% spolehlivostí „calls“ reference nebo zjištěných variant, a některé mezery dosekvenovává Sangerovým nebo srovnatelným sekvenováním. Tento typ testu je možné použít **pro potvrzení diagnózy**, ale pro



její vyloučení nikoli. V testu typu B budou klíčové geny komplexně pokryty, jak bylo uvedeno v kapitole 2.2.2

3. **Test typu C:** se spoléhá pouze na kvalitu NGS, žádné dodatečné Sangerovo (nebo jiné) sekvenování není nabízeno. Na příklad, pokud je panel genů vybrán na základě dat z exomového sekvenování, bez doplnění analýzy dalším sekvenováním. Test typu C nebude splňovat kritéria „panelu klíčových genů“, ale laboratoř přesto musí specifikovat hodnotitelný rozsah NGS.

Doplněním MLPA nebo jiné nezávislé metody lze dále zvýšit citlivost testu, ale to patří do „diagnostické cesty/směrování“ spíše než do hodnotícího systému, který je míněn pro NGS a sekvenování.

**Prohlášení 2.05: Porovnání nabídek diagnostického testování mezi laboratořemi by měl umožnit jednoduchý hodnotící systém na základě pokrytí a diagnostického výtěžku.**

Navrhujeme, aby laboratoře informovaly o typu prováděného testu ve svých zprávách a na webových stránkách, aby mohli pacienti a indikující lékaři tyto testy a jejich ceny porovnat mezi jednotlivými laboratořemi.

EuroGentest tvoří databázi NGS panelů, která bude pravděpodobně dostupná na Orphanetu. Databáze by mohla hodnotící systém převzít pro zjednodušení porovnávání laboratoří

### 2.3 Srovnání s dalšími doporučeními

- nepřeloženo

## Kapitola 3: Informovaný souhlas a informace pro pacienta a lékaře

### 3.1 Úvod

Informovaný souhlas je základ genetického testování (Sequeiros et al., 2010), stejně tak konzultace před testováním. Jeho zavedení je doporučeno v zemích, kde není nutný ze zákona, z důvodu možných nevyžádaných nálezů. V zemích, kde je souhlas povinný, musí být upraven pro NGS analýzu.

### 3.2. Stanoviska a příklady

#### 3.2.1. Důsledky různých NGS testů

Dopady diagnostického testování založeného na NGS závisí na postupu, platformě, filtrovacích procesech a ukládání dat v laboratoři. Indikující lékař je zodpovědný za předání informace o testu pacientovi, proto by měl být plně informovaný o omezeních a možných neočekávaných dopadech genetického testování. Pokud je nabízen panel genů, pak je vyžadována znalost genů zahrnutých v tomto panelu.

**Prohlášení 3.01: Laboratoř musí pro každý NGS test uvádět: onemocnění, na které/která je test zaměřen; názvy testovaných genů; jejich vykazovaný rozsah; analytickou citlivost a specifičnost testu; a pokud je to možné i onemocnění, která nejsou ke klinickému fenotypu relevantní, která by ale mohla být způsobena mutacemi v testovaných genech.**

Důsledky testu založeného na NGS vychází především z možnosti výskytu nevyžádaných a sekundárních nálezů. Nevyžádané nálezy jsou zachyceny v genech spojených s testovaným onemocněním, sekundární nálezy jsou detekovány v genech, které s etiologií testovaného onemocnění nesouvisí.



**Prohlášení 3.02:** Aby se snížila možnost výskytu sekundárních nálezů, měly by se diagnostické laboratoře ve svých analytických postupech soustředit na to, jaké geny jsou součástí panelu a na jejich odpovídající validaci.

Možnost nevyžádaných nálezů v panelu genů je velmi nízká a je závislá hlavně na zařazených genech. Některé geny, či specifické sekvenční varianty, v panelu genů mohou souviset s onemocněními, která neodpovídají klinickému fenotypu. V takových případech se nelze sekundárním nálezům vyhnout. **U recesivních onemocnění** mohou být navíc detekovány **heterozygotní mutace, čímž dojde k odhalení přenašeče onemocnění**. To bude mít samozřejmě důsledek pro klinicko genetické poradenství. Tato informace by měla být uvedena ve zprávě (viz Kapitola 5).

**Prohlášení 3.03:** Laboratoře by měly poskytovat informace o možnosti nevyžádaných nálezů.

Laboratoře by měly poskytnout obecné prohlášení o tom, že výsledky z analýzy panelu genů mohou odhalit širší fenotyp, než který se týká původní diagnózy.

V každém případě by měl klinik před indikací NGS testu zvážit a zkontrolovat řadu parametrů:

1. technické aspekty, tj. být si vědom toho, že jde o rozsáhlý test ve srovnání s analýzou jednoho genu, citlivost testu může být omezená v závislosti na diagnóze;
2. riziko nevyžádaných a sekundárních nálezů u konkrétního NGS testu;
3. diagnostickou indikaci, tj. musí být předepsán odpovídající test;
4. pro správnou interpretaci výsledků a pro napsání adekvátní výsledkové zprávy je důležité, aby měla laboratoř o pacientovi klinické informace. Klinik by měl tyto informace laboratoři poskytnout.

Pokud si lékař není některým z těchto bodů jistý, měl by požádat o radu nebo od indikace NGS testu ustoupit. Lékař musí mít kontakt na osobu, která je za NGS test zodpovědná. Laboratoř by měla poskytnout informace o odpovědných osobách, na které se může lékař v případě potřeby obrátit.

Kvalita interpretace nevyžádaných a sekundárních výsledků (ve smyslu patogenní (s dopadem na funkci proteinu) vs. neutrální vs. „VUS“) by měla být stejná jako pro zbytek testu.

### 3.2.2 Postup pro šíření nevyžádaných a sekundárních nálezů

Před realizací NGS testu je v klinickém (genetickém) centru nutné nastavit "**postup pro nevyžádané a sekundární nálezy**", který by měl být v souladu s rozhodnutími etické komise. Mělo by být rozhodnuto, zda jsou pacientům kromě původního diagnostického výsledku nabízeny **opt-in, nebo opt-out možnosti**. Jsou-li tyto možnosti poskytovány, měli by být na základě závažnosti onemocnění, věku jeho počátku, mortality, existence efektivní léčby atd., klasifikovány různé výsledky. Použitelné klasifikační modely již byly publikovány (Berg et al. 2011, Bredenoord et al. 2011), ale možnosti, které lze nabídnout velmi závisí na místních zásadách. Postup by měl také určovat, zda jsou nevyžádané nálezy a **status přenašeče interpretovány ve zprávě**.

**Prohlášení 3.04:** Pokud se klinické centrum nebo laboratoř rozhodne pacientům podávat informaci o přenašečství u nesouvisejícího onemocnění a o sekundárních nálezech, měly by zavést tzv. **opt-in/opt-out postupy s veškerými logistickými náležitostmi**.

V hlavní zprávě by měly být hlášeny nevyžádané nálezy a status přenašeče u genů zahrnutých v panelu. Sekundární nálezy by měly být pospány v samostatné zprávě. Sestavení multidisciplinárního výboru odborníků nebo místní etické komise k účelu diskuze o odeslání sporného sekundárního nálezu indikujícímu lékaři je volitelné.

Pokud není dostupná etická komise, např. v případech, že se jedná o komerční laboratoře nabízející NGS testování v klinickém kontextu, měla by probíhat pravidelná panelová diskuze odborníků na téma, jak se vypořádat s nevyžádanými nálezy a zda jsou výsledky soudně napadnutelné či nikoliv. Panel by měl být složen přinejmenším ze tří odborníků s klinickými zkušenostmi včetně atestovaných klinických genetiků a lékařů jiných specializací, kteří se přímo podílejí na péči o pacienta. Jednotlivé případy a výsledky diskuzí by měli být dokumentovány ve formuláři, který účastníci diskuze podepíší.

### 3.2.3 Poradenství při použití NGS diagnostických testů

Je nezbytné zajistit pacientovi genetické poradenství již před provedením testu, kde by mu měly být vysvětleny očekávané výsledky (negativní, pozitivní) a pacient by měl být upozorněn na možnost nevyžádaných a sekundárních nálezů. Měly by být definovány oba typy nálezů a pravidla laboratoře týkající se šíření těchto nálezů..

**Prohlášení 3.05:** Pacientovi by měla být jasná pravidla o poskytování nevyžádaných a sekundárních nálezů dané laboratoře.

Měla by být poskytnuta informace o interpretaci výsledků, zvláště fakt, že se tyto informace mohou změnit se vzrůstajícími znalostmi. Pojem nevyžádané a sekundární nálezy musí být projednány ve fázi před provedením testu. Doporučujeme písemný informovaný souhlas, který ale není vyžadován, pokud si lze vybrat z několika možností vrácení nevyžádaných a sekundárních nálezů.

**Prohlášení 3.06:** Doporučuje se poskytnout pacientům informace písemně formou letáku, nebo online na webových stránkách.

Souhlas musí zahrnovat oddíl o sdílení anonymizovaných variant v populačních a s onemocněním souvisejících databázích. Musí být respektovány zásady ochrany osobních údajů a zákony v jednotlivých zemích. V klinické praxi by měla být podporována účast při sdílení variant. Umožní to snadnější interpretaci variant a přinese to prospěch ostatním pacientům. Pokud záchyty v panelu genů exomu nebo genomu nevyřeší diagnózu, měla by být provedena druhá celoxomová nebo celogenomová analýza. Během tohoto poradenství by měl být vyplněn nový informovaný souhlas.

### 3.3. Srovnání s jinými doporučeními

-nepřeloženo

## Kapitola 4: Validace testu

### 4.1 Úvod

U diagnostiky pomocí NGS musí být empiricky stanoveny a validovány přesnost, analytická preciznost, analytická citlivost a specifita, vykazovaný rozsah výsledků testu a referenční rozsah.

#### 4.1.1 Definice

Platforma nezahrnuje jen NGS sekvenátor, ale také izolaci DNA, obohacovací metody, přípravu knihoven a analýzu dat. Účelem validace platformy je stanovení, že masivní paralelní sekvenační systém je schopen správně přečíst sekvenci DNA (Gargis et al. 2012). Zároveň umožní vyhodnotit, jak přesně může být detekována každá varianta. Součástí validace není interpretace nalezených variant.

#### 4.1.2 Popis analytického postupu

V tomto dokumentu jsou doporučení pouze pro analýzy germinálních mutací (ne *de-novo* ani somatické mutace). **Kvalita vzorku** vzniká kombinací mnoha parametrů, jako je **množství vytvořených dat, podíl PCR duplikátů a pokrytí (coverage)**. V diagnostickém prostředí, musí být analyzované pouze vzorky dobré kvality. Je proto nezbytné definovat kritéria, která charakterizují vysokou kvalitu **cílených genových panelů**, exomů nebo genomů.

Obecně platí, že postup analýzy NGS dat je složen z detekce bází (base calling) a demultiplexingu, mapování, anotace a filtračních kroků. Tyto kroky jsou popsány v tabulce 1, včetně dalších volitelných kroků.

#### 4.1.3 Parametry kvality

V diagnostickém nastavení musí být posouzeny pouze vzorky v dobré kvalitě. Je proto nezbytné definovat kritéria pro charakterizování vysoce kvalitních cílených genových panelů, exomů nebo genomů.

Kvalita vzorku by měly být hodnocena na třech úrovních:

- **Technický cíl:** vyhodnocovat kvalitu vzorku na základě technického cíle je férové v případě capture sekvenování, v případě exomového sekvenování je to závislé na použitém kitu.
- **Klinický cíl – Region of interest (ROI):** klinický cíl (geny, úseky, které budou interpretovány) musí být zváženo předtím, než se nastaví vykazovaný rozsah a navrhne diagnostický test (viz kapitola 2 a následující oddíl). Kvalita vzorku nesmí být posuzována pouze podle ROI, musí být posouzena i z technického hlediska.
- **Seznam transkriptů:** laboratoře se navzájem liší použitými kity, panely genů i klinickými cíli (budou se tedy lišit jejich seznamy transkriptů). Z důvodu porovnání kvality mezi laboratořemi by měla být kritéria kvality vypočtena nejen na základě seznamu transkriptů, ale na základě všech kódujících transkriptů z referenční sekvence.

Ačkoliv hraje cíl při hodnocení kvality vzorku významnou roli, kvalita není závislá pouze na cíli, jde o kombinaci mnoha parametrů. Je nutné vzít v úvahu také množství vytvořených dat, podíl klusterů přiřazených ke každému vzorku (při multiplexování), podíl PCR duplikátů a pokrytí (coverage). Stejně tak nestačí pokrytí samo o sobě, zejména pokud bereme v úvahu hrubé pokrytí. Kritéria kvality by měla být založena na informativním pokrytí namísto hrubého pokrytí (Weiss, Van der Zwaag et al. 2013). Geny, u nichž se vyskytují pseudogeny nebo repetitivní sekvence mohou vykazovat vysokou úroveň hrubého pokrytí, ale nízkou úroveň pokrytí informativního (pokud jsou vyřazena všechna mapována čtení se špatnou kvalitou).

Podíl cíle, u něhož může být spolehlivě stanoven genotyp, tj, pro který je získáno dostatečně **informativní pokrytí** pro přesné zjištění genotypu, poskytuje stručné hodnocení kvality, které lze aplikovat na tři výše definované cíle. Pokud se podaří všechny kroky přípravy vzorku, mělo by být toto číslo vysoké a reprodukovatelné. Nicméně pokud jeden krok selhal, měl by být podíl cíle, u něhož lze spolehlivě stanovit genotyp, nižší. Přítomnost mnoha PCR duplikátů v důsledku selhání přípravy knihovny může například snížit celkovou coverage a snížit tak počet míst se spolehlivě stanovitelným genotypem. Také nízké množství dat může vést k nízké informativní coverage.

**Prohlášení 4.01:** Všechna hodnocení kvality používaná při NGS diagnostických postupech by měla být přesně popsána.

Dobře zdokumentovány by měli být zejména detaily výpočtu měření, aby byla interpretace měření jasná. Pro usnadnění automatické manipulace s hodnotami řízení jakosti (QC), by měla být kvalita měření definována a dokumentována v ujednocené terminologii a měl by být použit standardizovaný formát. Např. qcML projekt udržuje obecný souborový formát XML pro uložení QC dat a ontologii QC termínů pro genomiku a proteomiku.

Technologie NGS tedy vyžaduje monitorování specifit běhu analýzy/specifických rysů vzorků. Údaje z monitorování nemusejí být hlášeny, ale měl by být použit pro průběžnou validaci.

#### **4.1.4 Monitoring a sledování vzorků**

Technologie NGS vyžaduje monitorování **specifických vlastností běhu** jako je počet poolovaných vzorků, podíl klastrů přiřazených ke každému vzorku a skóre kvality podle pozice báze. Každý sekvenační běh musí být sledován, zda jsou nebo nejsou splněny specifikace přístroje. Kromě toho by měly být vymezeny minimální požadavky pro důležitá kvalitativní měření (tj. kvalita báze, délka čtení, atd. podle charakteristiky platformy). Měla by být sledována také **analýza specifických rysů vzorku**, jako je informativní pokrytí, uniformita pokrytí, bias v rámci řetězce, GC bias, kvalita mapování, podíl mapovaných readů, podíl duplikátů, podíl cíle pokrytého minimální hloubkou coverage, podílem nepokrytého cíle, průměrnou coverage, přesnosti stanovení genotypu (calling), počet variant, a poměr transice/ transverze. Některá QC měření, která by měla být rutinně sledována pro všechny vzorky, jsou uvedeny detailněji v tabulce č. 1

**Table 1: Elements of a NGS bioinformatics pipeline**

Processing step	Description	Tools and databases	Output
<b>Base calling and demultiplexing</b>	Base calling and demultiplexing, are also referred as primary analysis.	vendor software of the sequencing platform	FASTQ file(s)
<b>Primer trimming</b>	In amplicon sequencing primers have to be trimmed from the reads	[CutAdapt (Martin et al. 2011)], [BWA (Li & Durbin, 2009)] (soft clipping while mapping)	FASTQ files or BAM file (if soft clipping by a mapper such as BWA)
<b>Adapter trimming (optional)</b>	Sequencing adapters may be trimmed from the read ends for those reads where the insert size is smaller than the read length. If not trimmed, sequenced adapters may interfere with mapping and variant calling, leading to false-positive or false-negative variant.	[Trimmomatic (Bolger et al. 2014)], [SeqPrep ( <a href="https://github.com/jstjohn/SeqPrep">https://github.com/jstjohn/SeqPrep</a> )], [CutAdapt (Martin et al. 2011)], [BWA (Li & Durbin, 2009)] (soft clipping while mapping)	FASTQ files or BAM file (if soft clipping by a mapper such as BWA)
<b>Low-quality trimming (optional)</b>	Low quality bases may also interfere with mapping and variant calling and can be trimmed from the end (and begin) of reads.	[Trimmomatic (Bolger et al. 2014)], [SeqPrep ( <a href="https://github.com/jstjohn/SeqPrep">https://github.com/jstjohn/SeqPrep</a> )], [CutAdapt (Martin et al. 2011)], [BWA (Li & Durbin, 2009)] (soft clipping while mapping)	FASTQ files or BAM file (if soft clipping by a mapper such as BWA)
<b>Mapping</b>	In the read mapping step, paired-end/single-end reads are mapped to the reference genome allowing for base changes and indels. Mapping should always be performed against the full reference genome even when a small gene panel is sequenced. .	[BWA (Li & Durbin, 2009)], [Noalign ( <a href="http://www.novocraft.com/main/index.php">http://www.novocraft.com/main/index.php</a> )], [Stampy (Lunter & Goodson 2011)], [SOAP2 (Li et al. 2009)], [LifeScope – for color space reads ( <a href="http://www.lifetechhnologies.com">http://www.lifetechhnologies.com</a> )], [Bowtie (Langmead & Salzberg 2012)]	BAM file

<b>Duplicate removal (optional)</b>	In shotgun sequencing few duplicates are expected since the DNA is randomly sheared. However, duplicates can occur during PCR and as an artifact of imaging. In amplicon sequencing, duplicates are expected and should not be removed.	[Picard MarkDuplicates ( <a href="http://broadinstitute.github.io/picard">http://broadinstitute.github.io/picard</a> )]	BAM file
<b>Indel realignment (optional)</b>	The presence of indels in the sequenced samples often leads to multiple single base mismatches around these sites, especially if they reside close to the start or end of reads. These artifacts may show up as false-positive variants during subsequent analysis. Local re-alignment algorithms identify such positions and try to minimize the amount of mismatching bases by performing a local re-alignment of the indel spanning reads, increasing the accuracy of the calls while minimizing false positives.	[GATK RealignerTargetCreator & IndelRealigner (DePristo et al. 2011)] and [SRMA (Homer & Nelson 2010)]	BAM file
<b>Quality score recalibration (optional)</b>	After mapping to the reference genome, the base quality score of the reads can be recalibrated to better match the probability of false base calls and to spread the quality scores wider over the valid range. In most algorithms, false base calls are distinguished from real variants by performing a simple base calling or using databases of known polymorphisms, e.g. [dbSNP].	[GATK BaseRecalibrator & PrintReads (DePristo et al. 2011)], [ReQON (Cabanski et al. 2012)]	BAM file
<b>Variant calling</b>	Variant calling consists of detecting and genotyping differences to the reference genome (base changes and small indels).	[samtools (Li et al. 2009)], [GATK UnifiedGenotyper (DePristo et al. 2011)], [GATK HaplotypeCaller (DePristo et al. 2011)] and [Platypus (Rimmer et al. 2014)]	VCF file
<b>Annotation</b>	Variant interpretation requires detailed annotation. Very basic annotations are gene name, region (exonic, splicing, intronic, intergenic, etc.) and coding change information. Additionally, minor allele frequency for known polymorphisms, pathogenicity and conservation scores and clinical databases can be used.	[AnnoVar (Wang et al. 2010)], [SNPeff (Cingolani et al. 2012)], [Cartagenia Bench Lab NGS ( <a href="http://www.cartagenia.com/products/bench-lab-ngs/">http://www.cartagenia.com/products/bench-lab-ngs/</a> )] [dbSNP (Sherry et al. 2001)], [1000 Genomes (The 1000	CSV, TSV, TXT, excel files or databases



		<p>Genomes Project Consortium 2012]], [ESP 6500 (<a href="https://esp.gs.washington.edu/drupal/">https://esp.gs.washington.edu/drupal/</a>)]</p> <p>[SIFT (Kumar et al. 2009)], [PhyloP (Cooper et al. 2005)], [MutationTaster (Schwarz et al. 2010)]</p> <p>[COSMIC (Forbes et al. 2008)], [OMIM (<a href="http://omim.org/">http://omim.org/</a>)], [ClinVar (Landrum et al. 2014)], [HGMD (Stenson et al. 2014)]</p>	
<b>Filtering</b>	<p>To find disease related variants in large variant lists, rigorous filtering is needed. Typical variant filters exclude low quality variants, intronic/intergenic variants, synonymous SNPs or known polymorphisms with low frequencies in the population. However, this kind of filtering selects both for deleterious and false-positive variant calls. To remove the false-positives, filtering according to variant frequencies of an <i>in-house</i> database, containing all the processed samples of a lab, is often applied. Because an <i>in-house</i> database accumulates false-positive variants that are specific for the used sequencing platform, sequencer and analysis pipeline, it can be used to identify and remove these false-positives.</p>	<p>[SnpSift (Cingolani et al. 2012)], [Cartagenia Bench Lab NGS (<a href="http://www.cartagenia.com/products/bench-lab-ngs/">http://www.cartagenia.com/products/bench-lab-ngs/</a>)]</p>	<p>CSV, TSV, TXT, excel files or databases</p>



**Prohlášení 4.02:** Diagnostické laboratoře musí zavést strukturovanou databázi relevantních měření kvality v rámci (i) všech typů platform (ii) všech typů testů (iii) a všech zpracovaných vzorků.

**Údaje ze sledování by neměli být hlášeny, ale použity pro průběžnou validaci.**

Je důležité sledovat výjimky, jako jsou např. počty případů, kdy byl vzorek sekvenován, aby bylo dosaženo definovaného kritéria kvality a **korekce případné záměny vzorku**. Musí být použita metoda sledování vzorku, protože pracovní postupy při NGS analýzách jsou velmi komplexní a zahrnují více kroků zpracování a to jak v laboratoři, tak následně při výpočetní analýze. Například, obyčejná SNPs by mohla být považována za obohacené cíle a genotypovány nezávislými metodami (tj. sekvenováním nebo qPCR genotypováním, viz dodatek 2). **Pro vzorky, které byly zaměněny, a pro jejichž záměnu není vysvětlení, by neměla být vypracována zpráva.**

**Prohlášení 4.03:** Součástí nastavení celé metody musí být i jednoznačná identifikace vzorků a možnost sledování (z hlediska nezaměnitelnosti vzorků, v originále zmiňují využití čarových kódů) v celém průběhu jejich zpracování. Tyto aspekty musí být zahrnuty ve validaci platformy.

Při ověřování platformy, musí laboratoř zajistit, aby všechny její přístroje a činidla splňovala požadavky výrobců. Musí být zjištěna omezení každé technologie a brána na zřetel během vývoje testu a analýzy dat. Laboratoře mohou rozlišovat funkce (pro ověření), které patří k platformě, specifickému testu, nebo analytickému postupu. Měl by se sledovat také podíl nezmapovaných readů a nezařazených MIDů, protože to umožní nalézt poškozené vzorky (kvůli kontaminaci vzniklé v průběhu pracovního postupu). K SNPs, které se využívají k identifikaci vzorku, mohou být navíc přidány oblasti pro kontrolu kvality u všech panelových nebo exomových obohacení. Výpočet počtu zachycených aberantních bází (non-wt calls), zachycených neplatných bází (označených jako báze 'N') a sporadických indels v těchto regionech může pomoci identifikovat poškozené vzorky. Navíc srovnávání těchto parametrů umožňuje přímé porovnání různých verzí diagnostického testu a zároveň porovnání mezi testy. Mohou se porovnat různé sekvenační platformy, metody pro enrichment atd. a tyto oblasti by tak umožňovali zkoušet způsobilost. Varianty zjištěné v oblastech kontroly kvality musí být vyloučeny z metrických výpočtu. Doporučuje se použít 3 velké exony na různých chromozomech, které neobsahují známé polymorfismy, především insdely (Tabulka č. 2). Použití tří regionů místo jednoho, poskytuje zálohu v případě velkých delecí nebo problému s enrichmentem. Exony jsou použity, protože jsou již zařazeny do exomového obohacování a tak musí být přidány jako vlastní obsah k panelu.

*Table 2: Quality control regions*

chromosome	start (hg19)	end (hg19)
chr1	152057442	152060019
chr9	5919683	5923309
chr18	19995536	19997774

**Komentář k a priori šanci nalezení varianty - Příklady**

Představme si, že existuje 99% šance zachytit heterozygotní variantu při pokrytí 20x. Toto ovlivní frekvenci záchytu mutace asociované s onemocněním a to různě, v závislosti na různých přístupech, ale také v závislosti na dědičnosti onemocnění (z důvodu zjednodušení předpokládáme, že pokrytí nižší než 20x má 0% šanci zachytit heterozygotní variantu, což samozřejmě není úplně pravda):

U recesivních onemocnění

A. Pro exomové sekvenování

Pokud je 75% exonů analyzováno s pokrytím 20x,

- budou zachyceny dvě varianty složeného heterozygota v jednom genu jen v 55,1% případů
- ve 38,3% případů bude nalezena pouze jedna varianta;
- a v 6,6% nebude zachycena ani jedna varianta

Pokud je 86% exomu pokryto 20x

- budou zachyceny dvě varianty složeného heterozygota v jednom genu jen v 72,5% případů
- ve 25,3% případů bude nalezena pouze jedna varianta;
- a v 2,2% nebude zachycena ani jedna varianta

Pokud je v cílovém panelu 96% cíle pokryto 20x

- budou zachyceny obě varianty složeného heterozygota v jednom genu jen v 90,3% případů
- ve 9,5% případů bude nalezena pouze jedna varianta;
- a v 0,2% nebude zachycena ani jedna varianta

B. Pro celorexomové sekvenování,

Pokud je 75% exomu pokryto 20x

- bude 1 heterozygotní varianta 1 genu zachycena v 74,2% případů,
- u 25,8% případu nebude varianta nalezena

Pokud je 86% exomu pokryto 20x

- bude 1 heterozygotní varianta 1 genu zachycena v 85,1% případů,
- u 14,9% případu nebude varianta nalezena

Pokud je v cílovém panelu 96% cíle pokryto 20x

- bude 1 heterozygotní varianta 1 genu zachycena v 95% případů,
- u 5,0% případu nebude varianta nalezena

## 4.2 Stanoviska a příklady

### 4.2.1 Validace platformy

V průběhu validace platformy se laboratoř musí ujistit, že všechna její zařízení a reagentie splňují požadavky výrobců. Musí být zjištěna omezení každé technologie a vzata do úvahy při analýze dat a vývoji testu.

**Prohlášení 4.04:** Přesnost a preciznost (accuracy and precision) by měly být součástí obecné validace platformy. Validace nemusí být pro jednotlivé metody nebo testy opakována.

Každá sekvenační technologie má své silné a slabé stránky. Bioinformatické nástroje proto musí odrážet tyto charakteristiky. Musí být použity nejaktuálnější referenční sekvence. **Potřeba přiměřeného pokrytí je závislá na typu přítomné varianty a jejího počtu kopií** (homo, hemi, heterozygot). Tento parametr a prahové hodnoty pro procento alelického přečtení by proto měli být empiricky stanoveny a v průběhu validace testu také validovány. Ve srovnání s heterozygoty stačí pro přesnou detekci SNP u homozygotů či hemizygotů nižší coverage.

- **Přesnost** se odkazuje na souhlas opakovaných měření stejného materiálu. Ke stanovení přesnosti na základě posouzení **reprodukovatelnosti** (přesnost mezi sériemi - běhy) a **opakovatelností** (přesnost v sérii - rámci běhu) by měl být analyzován **adekvátní počet vzorků (min 3)**.

- **Opakovatelnost** lze stanovit přípravou a sekvenací stejného vzorku vícekrát (min 3x) za shodných podmínek a následného vyhodnocení vzájemného souladu v detekci varianty a výkonnosti.
- **Reprodukovatelnost** hodnotí konzistenci výsledků u stejného vzorku za různých podmínek, jako například analýza vzorku v různých bězích, rozdílná příprava vzorku, příprava vzorku různými pracovníky a použití rozdílných přístrojů. Shoda mezi 95 a 98% je považováno za uspokojivé.

#### 4.2.2 Validace analytického postupu

**Prohlášení 4.05:** Bioinformatická analýza musí být upravena pro konkrétní použitou technickou platformu.

Při ověřování používaného postupu musí být měřeny diagnostické specifikace na základě vyhodnocení analytické citlivosti a specifičnosti. Lze k tomu použít několik metod:

- Srovnání genotypů získaných z dg. testů s genotypy získanými z SNP arraí; nicméně toto srovnání může být zatíženo chybami, protože dbSNP varianty zahrnuté ve většině SNP arraích jsou obvykle používány k procvičení a posílení genotypovacích algoritmů.
- **Slepé srovnání genotypů získaných z diagnostického testu s variantami ověřenými sangerovým sekvenováním**, nevýhodou této metody bývá nízký počet běžně dostupných variant
- **Srovnání získaných genotypů použitím dvou rozdílných NGS technologií**
- Analýza umělého datového souboru, v němž jsou známy skutečné varianty a chyby
- Resekvenování a/nebo analýza dobře charakterizovaných veřejně dostupných vzorků DNA jako jsou 1000genomes DNA vzorky dostupné přes Coriell repositories, zatímco dopovídající sekvenční datové soubory jsou k dispozici na adrese [www.1000genomes.org](http://www.1000genomes.org)

Dostupnost **velice dobře charakterizovaných vzorků** je ideální situací. Směrem k tzv. „platinovým“ datovým souborům jsou již také vytvořeny přístupy (GenomeInABottle – <http://genomeinabottle.org>). Další projekt poskytuje otevřený přístup k datům, z vyčerpávajícím způsobem osekvenované třígenerační rodiny, jejíž **DNA vzorky mohou být objednány přes Coriell repository**. Jsou dostupné konsenzuální seznamy variant ze sekvenčních dat získaných ze tří technických platform, které byly plně validovány křížovými kontrolami nebo doplňkovými metodami. Vzorky DNA těchto osob může být použita pro validaci platformy nebo bioinformatického postupu. **V souladu s ověřovacími postupy stanovenými pro validaci Sangerova sekvenování (Mattocks et al . 2010), se doporučuje ověřit okolo 300 variant na platformu s cílem upřesnit senzitivitu a specifičnost systému.**

**Prohlášení 4.06:** Analytická citlivost a specifičnost musí být v průběhu validace analýzy stanoveny samostatně pro každý typ varianty.

Jakékoliv změny v chemii, obohacovacích protokolech, nebo platformě bioinformatické analýzy bude opravňovat k **revalidaci postupu**.

Obecně platí, že jsou laboratoře vyzývány k provádění zkoušek odborné způsobilosti, jakmile byl test validován a podílení se na programech externího hodnocení kvality jakmile budou k dispozici. Jedná se o požadavek normy ISO 15189 pro akreditaci zdravotnických laboratoří, ale je to také efektivní způsob sledování výkonnosti v laboratořích. Laboratoře jsou v tomto kontextu vyzývány ke sdílení dobře charakterizovaných vzorků a datových souborů pro spolupráci za účelem zlepšování a standardizace postupu v diagnostice.

**Prohlášení 4.07:** Diagnostická laboratoř musí validovat všechny části bioinformatického analytického postupu pomocí standardizovaných datových souborů kdykoli, kdy jsou implementovány relevantní změny – nové verze, aktualizace (platí pro veřejné online nástroje i komerční softwarové balíky).

**In-house databáze** zahrnující všechny relevantní varianty představují **důležitý nástroj** s cílem pro identifikaci artefaktů specifických pro danou platformu, sledování výsledků validace, a zajištění výměny proxy pro locus specifické databáze a meta-analýzy. Typicky, tato databáze by měla umožnit další anotace (například falešných pozitivit, zveřejněných mutací, segregujících variant, a tak dále), které výrazně zefektivňují diagnostický proces.

Při (automatické) klasifikaci variant je třeba dbát na výběr **cut-off (tj. frekvenci varianty v normální populaci)**. Cut-off se bude lišit v závislosti na očekávaném typu dědičnosti (dominantní, recesivní, X-vázané) a na databázi, která je používána jako reference.

**Prohlášení 4.08:** Diagnostická laboratoř musí používat strukturovanou databázi všech relevantních variant s aktuálními anotacemi.

**Data by se měla ukládat** ve standardních formátech FASTQ, BAM a VCF, které by měly být rovněž použity pro výměnu dat s jinými laboratořemi. Při ukládání výsledků analýzy, mají být kromě výsledků analýzy uloženy také soubory full-log. Log soubory by měly být pokud možno úplné, což umožní reprodukovat převod celé analýzy a FASTQ dat do diagnostické zprávy. Bohužel, stále neexistuje žádný (mezinárodní) konsensus ohledně toho, co by mělo být ukládáno. Nicméně, ukládání dat musí být **v souladu s národními požadavky a zdravým rozumem**. Pokud je uložen soubor BAM, musí z něj být umožněno generovat původní soubor FASTQ tj. měl by obsahovat nemapované ready a v případě, že byly ready oříznuty, musí být soubory FASTQ uloženy také. Uložené soubory VCF by měly obsahovat všechny varianty s dobrou kvalitou před filtrováním podle frekvence alely, pozice v genomu atd. Pokud jsou VCF soubory uloženy, je výhodné použít genomový VCF soubor (gVCF) (včetně informací ohledně pokrytých pozic) tak, aby z nich mohly být spolehlivě vypočteny frekvence variant. Je třeba se vyhnout vlastním formátům souborů dodavatelů, protože by mohly být obtížně čitelné, jakmile by prodejce přerušil používání tohoto formátu souboru. Aby byla zaručena integrita dat je doporučováno používat kontrolních součtů. Kromě výsledků analýzy musí být uloženy také kompletní log files. Tyto soubory by měly být pokud možno úplné, aby byla zajištěna reprodukovatelnost celého postupu od FASTQ souboru po výsledek. Log files by také měli obsahovat všechny nástroje a databáze spolu s použitou verzí nástroje a databáze/ časové razítko/ a parametrů. Postupy, nástroje a databáze by měly být archivovány. **Je doporučeno používat systém pro správu verzí.**

**Prohlášení 4.09:** Diagnostická laboratoř musí zařídit dlouhodobé uložení všech relevantních dat.

#### 4.2.3 Validace testu

Před validací musí být diagnostický test pečlivě optimalizován. Před spuštěním testu je důležité, aby byl **definované ROI nebo klinické cíle**, jako jsou např. všechny použité kódující oblasti a konzervovaná sestřihová místa (Ellard et al. 2012). Při popisu klinického cíle, musí být uveden název a verze použitého transkriptu. Klinický cíl musí být definován v souladu s evropskými pokyny pro osvědčené postupy pro geny a onemocnění, jako jsou genové karty (Dierking et al., 2013), genové dokumentace (<http://ukgtn.nhs.uk/find-a-test/gene-dossiers/>) nebo osvědčené postupy navržené EMQN dokumenty (<http://www.emqn.org/emqn/Best+Practice>). Vzhledem k tomu, že se seznam kauzálních genů neustále vyvíjí, musí být klinický cíl pravidelně aktualizován. Některé oblasti klinického cíle nemohou být spolehlivě sekvenovány a proto by měly být vyloučeny z vykazovaného

rozsahu. Klinicky relevantní oblasti, které nejsou zahrnuty do vykazovaného rozsahu (z technických důvodů), by měly být genotypovány jiným způsobem, jako je např. Sangerovo sekvenování. Při vývoji testu je třeba vzít v úvahu typy mutací, které mohou být detekovány, jakož i informace o prevalenci takovýchto mutací u konkrétního onemocnění.

**Prohlášení 4.10:** Během vývoje testu musí být definován „vykazovaný rozsah“, což je část „regions of interest“ (ROIs), pro kterou je spolehlivě generován určitý počet „calls“. Vykazovaný rozsah by měl být k dispozici lékařům (např. ve zprávě, nebo online).

Exomové sekvenování, které má za cíl dosažení vysokého diagnostického přínosu, nevyžaduje další doplňkové analýzy k dosažení vysokého pokrytí ve všech oblastech genomu, na které se vztahuje, ale potřebuje jasně sdělit lékaři, že test nemůže být použit pro vyloučení konkrétní klinické diagnózy (**vykazovaný rozsah**).

**Prohlášení 4.11:** Požadavky na „vykazovaný rozsah“ závisí na cíli testu.

**Při optimalizaci dg. testu je nutné stanovit cenu a TAT.** Všechny vzorky nespĺňující kritéria kvality (viz dříve) by neměly být použity pro rutinní diagnostiku.

**Výkonnost dg. testu** musí být hodnocena z hlediska přesnosti, analytické citlivosti, analytické specifičnosti a přesnosti. Vzhledem k tomu, že **přesnost** závisí na sekvenci a kontextu, bude se měnit v rámci genomu/exomu a **měla by být stanovena na úrovni testu, tedy pro každý ROI**. Analytická citlivost závisí na informačním pokrytí a vykazovaném rozsahu.

**Ve zprávě by měly být jasně uvedeny limity použitého testu.** Obvykle zahrnují repetitivní sekvence, pseudogeny, homologní oblasti, obsah GC párů, drop-outy alel, a skutečnost, že některé typy variant, jako transverze a inverze, nemohou být zachyceny a/nebo se neberou v úvahu pro dg. test (např. pokud nejsou z exomových dat odstraněny informace o CNV, ale mohlo tak být technicky učiněno). Nalezené varianty je nutno ověřovat, aby se zjistilo, zda nedošlo k záměně. Ověřené varianty lze **využít k validaci bioinformatického postupu**. Nicméně toto není nutno provádět v případě, že je použitá technologie rozsáhle validována. **Tímto způsobem lze určit oblasti se spolehlivě stanovitelným genotypem a konfirmovat pouze ty ležící mimo tyto oblasti.**

**Prohlášení 4.12:** Kdykoliv dojde k důležitým změnám v metodě, musí být zkontrolovány parametry kvality a reanalyzovány vzorky. Laboratoř by měla předem stanovit jaké vzorky a kolik jich bude reanalyzováno pokud k aktualizaci metody dojde.

Např. metoda by měla být re-validována, pokud je použit nový genom, při aktualizaci softwarového nástroje, při úpravě genového panelu (u cíleného re-sekvenování), při změně přístrojového vybavení nebo reagensí. Jakmile laboratoř test validuje, měla by se zapojit do zkoušek způsobilosti (EHK).

### 4.3 Srovnání s dalšími doporučeními

- Není přeloženo

## Kapitola 5: Psaní zpráv

### 5.1 Úvod

Při psaní zpráv je nutné vycházet z doporučení HGVS. Zpráva musí být psána srozumitelným a jasným způsobem, aby ji rozuměl i člověk, který není znalcem v oboru. Připojení seznamu s fenotypy maximalizuje kvalitu výsledku poskytovaného lékaři. Pro ujednocení bioinformatické analýzy a konzistenci dokumentace nalezených variant je výrazně doporučeno zahrnovat i genomové souřadnice. Exony není nutno uvádět, protože dochází často k aktualizacím referenčních sekvencí a počty exonů v genu se mohou časem měnit a vnášet nesrovnalosti. Výsledková zpráva by měla obsahovat identifikační údaje pacienta, shrnutí klinického stavu, specifikaci použitého genetického testu, výsledek a závěr.

## 5.2 Stanoviska a příklady

### 5.2.1 Minimální obsah zprávy

**Prohlášení 5.01:** Zpráva z NGS testování by měla na jedné stránce obsahovat identifikaci pacienta, diagnózu, stručný popis testu, shrnutí výsledků a hlavní nález.

Jedna strana tedy shrnuje všechny důležité údaje a další stránky mohou být jako doplňkové stránky s více detaily.

**Prohlášení 5.02:** Dříve, než laboratoř poskytne analýzu tohoto typu, měla by nastavit a zdokumentovat pravidla, podle kterých bude reportovat genomické varianty. Pravidla by měla být v souladu s mezinárodními doporučeními.

### 5.2.2 Klasifikace variant

Neliší se od sangerova sekvenování. Pěti-třídová klasifikace. Nepoužívat termín polymorfismus.

**Prohlášení 5.03:** Údaje o variantách nejasného významu (VUS, UV) musí být shromažďovány z důvodů případné definitivní klasifikace jejich významu.

### 5.2.3. Nevyžádané a sekundární nálezy

**Prohlášení 5.04:** Laboratoře by měly mít před spuštěním testu jasně definovaný postup/protokol pro případ řešení nevyžádaných a sekundárních nálezů.

### 5.2.4 Povinnost re-kontaktovat

**Prohlášení 5.05:** Neočekává se, že by laboratoř systematicky re-analyzovala stará data a reportovala nové nálezy, a to ani v případě, že dojde ke změnám v panelech klíčových genů.

**Prohlášení 5.06:** Laboratoř musí vytvořit vlastní databáze variant pro onemocnění, jejichž testování nabízí, aby byla schopna varianty vyhodnocovat.

## Kapitola 6. Rozdíl mezi výzkumem a diagnostikou

### 6.1 Úvod

S rostoucími možnostmi exomového a genomového sekvenování se rozdíl mezi diagnostickým a výzkumným testováním stírá.

Diagnostický test je jakýkoli lékařský test, který byl provedený z důvodu stanovení diagnózy či odhalení onemocnění. V genetice to znamená, že v genetickém materiálu jedince byla hledána pravděpodobně patogenní varianta, která by vysvětlovala pacientův fenotyp, nebo bylo účelem vyšetření zjistit míru rizika vývoje onemocnění, které se projevilo v rodině. Použití genetický test může být velmi specifický (např. konkrétní exon genu), ale i méně specifický (např. CNV). Diagnostický test je proveden ve specializovaných laboratořích, které se řídí normou ISO15189.

**Prohlášení 6.01: Diagnostický test je jakýkoliv test prováděný za účelem odpovědi na otázku související se zdravotním stavem pacienta.**

Výzkum je zpravidla cílen na odhalení a interpretaci nových faktů.

**Prohlášení 6.02: Výzkumný test je založen na hypotéze a výsledek může mít pro pacienta zapojeného do projektu omezený klinický význam.**

#### **6.2.2 Rozdíl mezi diagnostikou a výzkumem**

- opět vomáčekově popisují, že rozdíl mezi diagnostikou a výzkumem se rozměňuje.

**Prohlášení 6.03: Výsledky diagnostického testu mohou vést k hypotéze.**

#### **6.2.3 Jaký typ NGS může být proveden v diagnostické laboratoři?**

**Prohlášení 6.04: Diagnostické testy, jejichž primárním cílem je hledání diagnózy konkrétního pacienta, by měly být prováděny akreditovanými laboratořemi.**

#### **6.2.4 Povinnost potvrdit výsledek výzkumu v diagnostickém prostředí**

**Prohlášení 6.05: Před předložením výsledků výzkumu indikujícímu lékaři a/nebo pacientovi (klientovi), musí být tyto výsledky potvrzeny v akreditované laboratoři.**

#### **6.2.5 Sdílení mutací a variant v mezinárodních databázích**

**Prohlášení 6.06: Frekvence všech variant, které jsou detekovány sekvenováním vzorků zdravých jedinců v diagnostickém a/nebo výzkumném režimu, by měly být sdíleny.**

**Prohlášení 6.07: Všechny varianty uváděné ve výsledkových zprávách by měly být sdíleny v regionálních, národních i mezinárodních databázích.**