American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional Copy Number Variants

——美国医学遗传学会对出生后固有的拷贝数变异解读和报告的标准指南

# 研究背景

基因微阵列技术用于获得大量的拷贝数变异数据，已经成为评估产后个体智力障碍、自闭症或者其他先天性疾病的首要手段。当前，基因微阵列技术包括阵列比较型基因杂交法、单核苷酸多态性平台检测法，这些方法相对于传统方法，能够得到更高分辨率的全基因组覆盖的染色体不平衡重排检测。

拷贝数变异是是指大小在1Kb到数Mb范围的基因组之间DNA片段相对于参考基因组的不平衡重排。”CNV”这个术语并不能说明临床意义（区分良性CNV和致病CNV），也不能说明相对的剂量（有拷贝数缺失和拷贝数重复之分）。虽然许多常见的低拷贝重复序列已经被遗传学家很好地解读，但是众多良性CNVs、新发现的引起临床症状的致病CNVs等片段的解读仍然是很大的挑战。

受条件所限，当前的CNV解读和报告因不同的实验室而异，本文的目的是制定相关标准，确保对拷贝数变异的评估解读和报告标准的一致性。本指南的CNV检测虽然是针对产后的，但是和产前的CNV检测也有一定关系。另外，本文不涉及分析的有效性，不能应用于瘤形成突变的检查，不能全面覆盖所有的CNV类型（很多CNV十分特殊没有算法能够应用于将它们全部发现）。

# CNV系统评价和临床解读

## 熟悉已明确的邻接基因综合征(Contiguous gene syndrome)

只审查单个基因中大片段缺失/重复序列可能难以找到与综合征相关的缺失/重复片段。所以，在已知的临床上有显著表征的常见片段微缺失/微重复综合征中的相关低拷贝重复序列和关键区域需要和目标片段认真比对和鉴定。

## CNV 片段的大小

虽然CNV的大小和临床意义有关的认知有一定的合理性，但是事实上绝大多数的CNV还是良性的，具有临床显著症状的CNV毕竟是少数。所以，应该根据实验室的芯片特点和临床上相关合理性的条件设置CNV片段大小的阈值，而不是仅仅根据临床意义做出判断。

## CNV片段内的基因情况

首先，需要认真审查CNV片段特点，是特有的富含基因片段、没有相关基因、拥有重复序列片段或者是假基因。然后，研究基因内拷贝数重复或缺失引起的潜在表型效应时，需要考虑片段内的基因是否是剂量敏感基因(dosage sensitive gene)或者是致病基因(associated with clinical disorders)。需要注意的是，由于探针的覆盖存在间隙，所以因此得出的CNV断点(breakpoints)并不十分准确。

更进一步对CNV内基因情况的临床解读如下：

### 文献中已报道的致病突变基因

相关基因的解读可以围绕下面的几个方面展开，虽然所列内容并不完全但是已经可以很好地理解其成因机理。

1. 某基因由于单倍剂量不足(haploinsufficiency)而引起临床表型，在拷贝数增加(gains)的情况下可能不会出现临床症状。
2. 显性疾病往往是由于相关基因的功能获得性突变而不是剂量不平衡引起的，涉及这类基因的CNV可能不会有与之相关的临床症状也可能会出现完全不同的临床症状。
3. 只包含基因的一部分的拷贝数增加可能会破坏该基因或者改变相应的编码序列，同时当某基因被报道为单倍剂量不足，应该考虑是否为这种情况。
4. 隐性致病基因的单拷贝数缺失可能建议认定为携带者。
5. 只包含内含子片段的小片段CNV可能不会对基因功能造成影响。

### 文献中未报道的突变基因

尽量避免或者谨慎仅依据基因的预测功能，或者依据模式生物上的功能特征或者依据体外实验推断出的一个基因致病成分。这些推断仍然是猜测，除非在人群中被发现存在有显著的临床特征。

### 片段中不含有基因

如果没有相关的参考文献用于审查，可不报道该CNV片段。除非该片段的大小超出实验室事先设定的阈值或者与已知表型的基因区域接近且有明确的参照理由。

## 使用内部和外部的文献进行CNV片段比较

我们推荐各实验室利用自身积累的经验和相关文献数据库开展CNV片段的评估，并将CNV分类为良性CNV、致病性CNV和临床症状不确定的CNV。一个被注释得很好的CNV片段往往只需要周期性地调查是否有全新的成果与已知相矛盾，一般情况下，不需要进一步地审查和过多的异议；但一个CNV在此之前原实验室未研究过，就需要认真地参考全体人群(general population)的CNV数据库。因为在很多大型数据库中许多基因没有对应的表型，所以不能只是参考有临床表型的人群数据。病人相关数据和全体人群CNV数据库比较需要考虑以下几点：

### 在全体人群研究中的拷贝数不平衡的剂量

在全体人群中，被认定为良性的CNV专门指出是拷贝数增加的情况，如果在相同的区域发现CNV是拷贝数缺失的，就不能排除致病的可能性，反之亦然。

相似地，如果在全体人群中某CNV被发现主要是杂合缺失(heterozygous deletion)，如果在同一位点出现纯合缺失的CNV也不能排除其致病的可能性。

### 已报道的良性CNV和待测CNV的大小关系

进行大小比较的前提是两者中的基因序列内容相同。要注意的是，不同微阵列平台得到的同一个CNV的长度不尽相同，早期用BAC(bacterial artificial chromosome-based)微阵列芯片得到的良性CNVs大小普遍被估计过高了。

### 考虑患者和数据库中样本的性别

这一考虑对在男性中发现的X染色体连锁CNV的判定尤为关键，因为很多被认为是良性CNV的突变实际上是在女性携带者身上未表型。另外，如果数据库中的数据是排除性染色体后的分析，那么匹配上性染色体的CNV不具有代表性。

### 全体人群数据库的可靠性

在大群体研究中的大多数CNV都没有被实验验证，如果一个CNV只在一次研究或通过一个微阵列芯片出现须慎用。

### 考虑正常人群中的临床表征

在进行CNV判定时，应该考虑数据库中不同个体的临床症状的表现程度，因为基于事先确定的标准筛选出的正常人群，其中很有可能含有“真实的病人”。所以在将一个CNV划分为良性时需充分考虑相关基因的外显率、表现度、发病时间与双亲的印记效应等多种因素。CNV在人群中出现的频率相对较高，被文献报道多次确认过为良性，更有可能就是良性。当然被文献多次报道不是指相关文献数量较多，因为多数文献使用相似的参考文献，而是指在同一个体中多次的研究均能证实。

# CNV临床意义的分类

根据相关的临床意义，CNV片段主要被分成如下的三类：

## 致病

1. 即使外显率和表现度是变动的，但是如果某CNV被多篇有同行审查的文献指出有临床症状，说明该CNV为致病性的。
2. 大片段CNV虽然相关医学文献没有描述，但是包含被明确指出的致病基因片段会被认为是致病性的。
3. 虽然全部片段的临床症状不明确，但是CNV的致病机理是明确的，该片段被认为是致病的。
4. 例外：如果在已知的细胞遗传异质性区域发现明确的致病位点，需要慎重考虑，因为该区域存在大量可见的突变(>3-5Mb)。

## 良性

1. 某CNV被多篇有同行互评的文献或者有审查的数据库认为是良性的，尤其是明确机理或者在具有常见多态性(common polymorphism)，该CNV被认为是良性的。
2. 如果CNV要被认定为多态性，一般要在>1%的人群中存在，并且其作为良性突变的剂量要明确。存在某区域片段重复是良性的，但是片段缺失是致病性的。

## 临床意义不明确

这一部分的内容相当广泛，很多片段在后期的研究中会被明确是良性还是致病的，在未出现充分证据之前仍然被分为临床意义不明确。某CNV在被审核时，需要提供该CNV致病性或者良性的可能性以及支持相关判断的可靠证据，同时要保留该分类判断的不确定因素。临床意义不明确建议分成如下几种，以下分类没有包含所有情况。

### 可能致病性

可能有的情况如下：

1. CNV只在单一病例报告中被认定是致病的，但是有明确的断点和表型，并且与患者的临床表征之间有特异性和相关性。
2. CNV中含有与患者临床表征间有一定相关性和特异性的值得关注的基因功能，但是相关推论来自于单一模型的须慎用，基本上弃用。尤其是那些来没有明确特异性的指标（如智力障碍）或者和基因功能相关的证据有限（如仅有少量信息能表明某基因与神经元表达有关）的情况强烈建议弃用。

### 可能良性

1. CNV中未含有基因（当片段长度超过设定的阈值需要加以报道）。
2. CNV在全体人群的突变数据库中有少量的样本，但是没有表现出常见多态性(common polymorphism)。

### 其他

1. CNV中包含基因，但不确定该基因是否剂量敏感。
2. CNV在多篇文献或数据库之间的研究结论互相矛盾，且对于临床表征的坚实结论未建立。

# CNV报告的组成部分

以下内容是关于CNV报告如何详细描述一个CNV类型所需要的各部分内容，如何清晰地表达和临床意义相关的微阵列芯片结果，以及报告的其他必要组成内容。

## 报告标准

报告中需要包括相关数据请款和CNV筛选的标准。可以选择不报道良性CNV，尤其是那些常见的多态型，但是如果选择报道相关良性的CNV，那么对于它们的筛选标准应该和致病性CNV的要求一致。

## CNV的位置、大小和增减情况描述

相关描述使用人类细胞遗传学国际命名体制（ISCN）制定的术语，不为某些对ISCN不熟悉的临床人员替换掉相关清晰的专业描述。

1. 细胞遗传学位置：染色体号和染色体区带
2. 剂量（缺失或者重复）以及CNV的作用机理（单拷贝数缺失、多序列重复等），作用机理的评估还需要其他的测试方法。
3. CNV的大小和基因组中的坐标，尤其是那些CNV中所含基因不明确的片段，CNV最小和最大的坐标需要提供。

## CNV临床意义的描述

需要清晰地描述出每一个CNV的临床意义，并且在建议栏中被概括总结好，和支持相关解读的有效证据和文献。

## CNV内基因的标明

报告中需要指明CNV中含有的相关基因以增加报告的实用性。

1. 对于已知相关临床症状的片段，只需要提供对应综合征的名称和区间内临床相关的基因信息。
2. 对于临床症状不确定的片段，建议列出区间内所有的RefSeq基因，如果有条件可以优化有一定时间的相关医学文献内容。
3. 区间内的基因不建议提供该基因相关信息的网址链接，因为基因信息随时发生改变。
4. 如果报告中涉及的基因数量很少，建议将每个CNV中含有的基因进行整体性分析。

## 临床后续工作开展的建议

如果CNV被鉴定为致病性或者临床意义不明确，报告中需要包含相关遗传咨询的建议、CNV片段所有可能的遗传学表征和家族成员的评估。

如果CNV被鉴定为临床意义不明确，报告中还需要提及对该CNV最新医学文献的追踪，用于明确其临床意义。对于特定病人后续的临床跟踪需要依靠医生和病人之间不间断的联系。

# 额外临床信息的特殊考虑

## 隐性病症的携带者报告

一般认为，拷贝数缺失的发现将表明个体为隐性疾病的携带者。杂合隐性突变超出报告的实际用途通常建议不用报告，但是报告需要清楚说明隐性疾病携带者不会被告知，任何隐性疾病的临床风险需要告知。以下情况在选择不报告时需要慎重考虑：

1. 隐性疾病的症状明确、人群中携带率较高、筛查手段可行。相关的报告将会对产前咨询以及相关先证者和家庭成员的疾病诊断提供依据。一般情况下，有证据表明该突变是偶然性的可以选择不报告，另外也不需要向医生或患者承诺携带者身份在任何条件的检测中都能确认。
2. 隐性疾病的临床表征与患者临床症状相吻合。这种情况下，建议开展近一步分子水平的检查。当然，这种特殊情况必须基于该隐性疾病有着明确的临床症状和详细记录。报告中需要清晰地阐述相关疾病的隐性性状，在没有其他相关突变该CNV不能诊断为该疾病的致病原因。

## 就诊时未出现(未检测出)该CNV所引起的临床症状报告

部分CNV虽然和相关疾病没有直接关系，但是可以作为发病前（或症状为检测出前）明确的诊断依据。鉴于不可能预先定义好所有可能发病的症状，所以建议对相关CNV进行报告，便于病人提前就诊；如果选择保密需要在报告中事先说明。

## 与肿瘤形成高风险的CNV报告

相关已知或者推定的抑癌基因存在拷贝数缺失，需要认真考虑。如果该基因是致病的胚型突变，相关致病机理、外显率、存活期以及临床处理办法需要在报告中说明。有关推定的抑癌基因的相关推测尽量避免在报告中出现，尤其是在人体中缺乏明确的胚型突变或者只是在相关的模式生物中出现过的。

# 家系报告的特殊考虑

当CNV的临床表征不够明确时，需要确认该CNV是否来自遗传或者是de novo突变。临床表征的显著与否很大程度上与划分的标准有关，所以需要强调相关工作很有难度，实际中又经常会出现只根据单一家系的遗传方式轻易地划分临床表征。理论上，应该通过对包含大量的受指定CNV影响或者不受指定CNV影响的家族成员的多组家系，或者携带同一CNV的不同个体进行研究才能得到可靠的临床表型的归类依据。

由于相关数据不够充分，所以基于有限的家族信息得出的推论需要非常慎重。对基于家庭成员的推测信息的免责声明需要在原始报告的附录中写明。每个CNV和每个家系都需要有格外的考虑，相关方向的研究调查也需要实验人员和临床医生开展专门的沟通。

## De novo CNV

1. 如果CNV在先证者中被发现为de novo突变，通常有明确的证据能够推定其为致病性的CNV。父系的研究不需要也不推荐进行，除非有明确的临床证据并且获得相关许可。
2. 在易发生变异的区域，某写CNV可能确实是de novo突变，然而却没有相关临床症状。
3. 如果只有一个亲本能开展后续研究，并且在该亲本上未发现相关CNV，那么关于该CNV的临床表征不需要相关推论。

## 遗传性CNV

当CNV在亲本或者其他家族成员中发现时，很多注意事项需要考虑。基本上不存在根据单一家系的遗传方式就能得出推论。亲本或者其他家族成员中的携带者需要开展全面的和先证者出现的临床症状有关的医学评估分析。如果这些信息没有提供给实验室，那么就不应该报道相关CNV的临床表征，并建议开展亲本和患者临床表征的相关性分析。

### 父母亲是患者

通常，可以根据证据谨慎地判断该CNV为临床症状的病因。但是为排除偶然性，需要对其他家族成员开展检测以进一步确认该CNV是遗传的还是散发的。

### 父母亲不是患者

通常，可以根据证据判断该CNV不为相关临床症状的病因，而可能是良性的。但是在得出结论前仍需要考虑如下的特殊情况：

1. 不完全外显：该CNV是致病的，但是父母只是携带者而没有相应表型。
2. 表现度不同：父母亲可能存在轻微的临床症状，或者是临床症状迟发。
3. 父母受印记效应影响：该CNV可能只有在特殊性别时才会体现出异常。
4. 先证者存在第二个未被检测到的突变：先证者从亲本遗传得到相关CNV，但在CNV的区间中不存在突变基因，即该CNV可能为良性的，而存在另一个致病性的CNV未被检测出来。
5. 父母为嵌合体：该CNV突变并不是存在于父母的所有组织中，不会表现出临床症状。
6. CNV在先证者和父母身上的大小不一致：在传代时，该CNV会罕见地出现被进一步修饰的可能性。如果亲本进行了相关检测方法（如原位杂交荧光检测），这种情况不能排除。
7. X染色体连锁的CNV：男性患者有X染色体连锁的CNV，其母亲也许是无表型的该CNV携带者。X染色体失活的研究可能可以提供信息，但并不是所有的X连锁疾病对女性携带者都表现出偏差。也可以观察母系家族内的其他男性成员的表型。

# 延伸讨论

随着全基因组检测的发展，临床遗传学转件肯定准备好对未知的发现展开相关的解读研究，而且考虑的内容也将不仅限于临床表征，还会包括社会、种族和法律因素。拷贝数变异的相关临床解读是复杂的工作，也是当前医学界努力的方向。正如本文中提出的各种考虑因素所表明的，CNV的解读是没有统一的全面而可靠的模式或者算法。因此，我们建议对相关微阵列芯片的解读需要由经过专业训练和合格资质的人员进行。另外，不同实验室的研究方法和解读方式需要提供，优质的实验室对每组测试集都需要开展细胞遗传学和分子遗传学两方面的专业研究。