Title

Recent challenges to the psychiatric diagnostic nosology: a focus on the genetics and genomics of neurodevelopmental disorders

Journal

International Journal of Epidemiology, 43(2)

ISSN

0300-5771

Authors

Kim, Young Shin
State, Matthew W

Publication Date

2014-04-01

DOI

10.1093/ije/dyu037

Peer reviewed

International Journal of Epidemiology, 2014, 465–475
doi: 10.1093/ije/dyu037
Advance Access Publication Date: 11 March 2014
Original article

Original article

Recent challenges to the psychiatric diagnostic nosology: a focus on the genetics and genomics of neurodevelopmental disorders

Young Shin Kim[1,2,3] and Matthew W State[4]*

[1]Yale Child Study Center, Yale University School of Medicine, New Haven, CT, USA, [2]Nathan Kline Institute for Psychiatric Research, Orangeburg, New York, USA, [3]Department of Psychiatry, Yonsei University College of Medicine, Seoul, South Korea and [4]Department of Psychiatry, University of California, San Francisco, CA, USA

*Corresponding author. Langley Porter Psychiatric Institute and Hospital, Department of Psychiatry, University of California, San Francisco, 401 Parnassus Ave, San Francisco, CA 94143, USA. E-mail: Matthew.State@ucsf.edu

Accepted 29 January 2014

Abstract

Recent advances in the genetics of neurodevelopmental disorder (NDD) have demonstrated that rare mutations play a role not only in Mendelian syndromes, but in complex, common forms of NDDs as well. Strikingly, both common polymorphisms and rare variations in a single gene or genetic locus have been found to carry risk for conditions previously considered to be clinically and aetiologically distinct. Recent developments in the methods and tools available for studying complex NDDs have led to systematic and reliable genome-wide variant discovery. Both common as well as rare, and structural as well as sequence, genetic variations have been identified as contributing to NDDs. There are multiple examples in which the identical variant had been found to contribute to a wide range of formerly distinct diagnoses, including autism, schizophrenia, epilepsy, intellectual disability and language disorders. These include variations in chromosomal structure at 16p11.2, rare *de novo* point mutations at the gene *SCN2A*, and common single nucleotide polymorphisms (SNPs) mapping near loci encoding the genes *ITIH3, AS3MT, CACNA1C* and *CACNB2*. These selected examples point to the challenges to current diagnostic approaches. Widely used categorical schema have been adequate to provide an entré into molecular mechanisms of NDDs, but there is a need to develop an alternative, more biologically-relevant nosology.

Thus recent advances in gene discovery in the area of NDDs are leading to a re-conceptualization of diagnostic boundaries. Findings suggest that epidemiological samples may provide important new insights into the genetics and diagnosis of NDDs and that other areas of medicine may provide useful models for developing a new diagnostic nosology, one that simultaneously integrates categorical diagnoses, biomarkers and dimensional variables.

Key words: nosology, genetics, genomics and etiology, neurodevelopmental disorders

Key Messages

- Recent advances in the genomics of common Neurodevelopmental Disorders (NDDs) have demonstrated that the identical genetic variant may increase the risk for a wide range of diagnoses formerly thought of as distinct. These findings are contributing to an ongoing re-conceptualization of the current psychiatric nosology. The use of epidemiological samples, studies grouping individuals based first on genetic findings, and efforts at combining existing categorical schema with dimensional phenotypes and biomarkers all promise to provide important new insights into the etiology and classification of NDDs.

Introduction

Recent advances in genetics research have presented fundamental challenges to long-held conceptions regarding diagnostic approaches to psychiatric syndromes. Nowhere has this process been more apparent than with regard to the genetics and genomics of neurodevelopmental disorders (NDDs), where recent evidence from a wide range of studies have pointed to indistinguishable risk factors for conditions that have long been held clinically and aetiologically distinct. Indeed, the very notion of specificity, which has figured so prominently in the conceptualization of diagnostic categorization, finds little support in the empirical data. Below, we address the current thinking regarding the nosology of NDDs and review the evidence from selected genetic studies that is contributing to a re-conceptualization of the boundaries between conditions. We also consider study designs, including the ascertainment of epidemiological samples, that will contribute to further progress in this area.

The Nosology of NDDs

NDDs are a heterogeneous group of clinical syndromes in which neurobiological development is disrupted. These disruptions lead to developmental delays or developmental deviations, which, irrespective of cause, restrict one or more areas of major life functions/activities.[1] It is generally assumed that the substrate for each of the NDDs appears early in development—as early as fetal development and not later than early childhood—with full expression of the related syndrome appearing at various points in the course of development.

The mechanisms underlying NDDs are often unknown but, when known, are as varied as the conditions themselves. Aetiological substrates cover the full spectrum of factors, ranging from environmental disruptions to genetically determined syndromes. In the final analysis, NDDs are highly prevalent conditions that have profound and long-lasting impacts on affected individuals, their families and communities. Because of their heterogeneity, NDDs pose particular challenges to classification and nosology that are fundamental to identifying like groups for the purposes of both research and treatment.

Depending on how a list is assembled, NDDs cover a very broad spectrum of clinical presentations that can include motor problems (including cerebral palsy), stereotypical behaviours [e.g. tics and obsessive compulsive disorder (OCD)], cognitive deficits [including mental retardation (MR)—now intellectual disability (ID)], communication deficits (CDs—expressive, receptive and social), learning disabilities (LDs), autism spectrum disorder (ASD), sensory deficits (hearing and/or vision), attention-deficit/hyperactivity disorder (ADHD) and epilepsy.

However, the already long list of commonly-accepted list of NDDs is not comprehensive enough. Indeed, there is a compelling argument to be made that the majority of currently classified neuropsychiatric syndromes are, in fact, NDDs. This criterion clearly applies to conditions that, in the current diagnostic nosology, must have onset before age 18 years, such as ASD, ADHD and Tourette disorder.[3] There is also room for debate about the appropriate classification of psychiatric conditions that are commonly thought of as 'adult' disorders, but, nonetheless, may show symptom onset in adolescence or even in childhood, including schizophrenia and mood and anxiety disorders. Even disorders that are thought to have a primary aetiological contribution from an environmental stressor, such as post-traumatic stress disorder (PTSD), also may usefully be thought of as a neurodevelopmental syndrome. This is based on the observation that not everyone exposed to a significant environmental event develops the same syndrome or, in fact, manifests clinically meaningful difficulties at all. One possible explanation for this heterogeneity in outcome is variations in neurodevelopmental vulnerability and resilience, thus placing the condition in a category with other, more traditional NDDs.

In the current psychiatric nosology, the vast majority of individuals with NDDs are classified solely on the basis of observable behavioural and/or morphological features. However, it is also important to note that the identification of the molecular substrates for certain of these conditions

has long been under way. Multiple NDDs have already been characterized with a specific genetic aetiology, especially those syndromes that include ID and/or epilepsy. Indeed, a review of the Online Mendelian Inheritance in Man (OMIM: http://www.ncbi.nlm.nih.gov/omim) database currently lists 152 NDDs for which specific causal genetic variants have been discovered.[4] A key point of distinction, however, is that gene discovery, until very recently, has largely been restricted to rare NDD subtypes that demonstrate Mendelian patterns of inheritance. So, although the number of distinct syndromes is large in total, this list accounts for a small fraction of the overall burden of morbidity from NDDs in the population.

To be sure, a rough dichotomy has emerged in the literature that either explicitly, or implicitly, distinguishes rare Mendelian subtypes from more common NDDs. This latter group of disorders is generally hypothesized to have heterogeneous and polygenic inheritance, involve a significant contribution from non-genetic factors and reflect a general absence of pathognomonic physical findings. The presumptions regarding the genetic architecture of common NDDs are based, in large part, on their prevalence, evidence for high heritability, the failure despite ample opportunities to identify a single or small number of genetic loci explaining a large proportion of risk, and a relative paucity of pedigrees showing simple Mendelian inheritance.

However, there are important exceptions to this informal classification scheme that separates 'idiopathic' or common disorders, from 'syndromic' or Mendelian NDDs: mutations in genes that are known to cause Mendelian neurodevelopmental syndromes may clearly present a clinical picture that is indistinguishable from common NDDs. Conversely, there have been recent examples of outliers among the population of individuals manifesting 'garden variety' NDDs that show Mendelian patterns of inheritance. For example, ASD is typically considered to be a common, genetically complex NDD. However, there are many molecularly-defined syndromes for which this behavioural constellation is prominent;[5] some of these syndromic forms of the disorder may present a clinical picture that, for all practical purposes, is indistinguishable from 'non-syndromic' ASD; and several recent successful efforts at gene discovery in 'typical ASD' have mapped genes in rare pedigrees that show Mendelian inheritance.[6–8] Nonetheless, for the sake of the present discussion, we will focus on the group of 'common, idiopathic NDDs,' characterized by: (i) relatively high prevalence (~1%); (ii) complex inheritance in the majority of cases; and, (iii) the absence of characteristic dysmorphology on physical examination.

These common conditions are enumerated in a section entitled 'neurodevelopmental disorders' in the recently published American Psychiatric Association (APA) Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-V), which relies on a syndromic/categorical diagnostic classification scheme. Among the NDDs defined in DSM-V are: Intellectual Development Disorder (formerly MR), CD [e.g. language disorder, speech disorder, fluency disorders (stuttering) and pragmatic disorder), ASD, ADHD, Specific LD (e.g. reading, writing, mathematics), Motor Disorders (e.g. coordination, stereotypies, tics) and Other NDDs (e.g. fetal alcohol syndrome)[9]. For a variety of reasons, other common syndromes, such as schizophrenia and OCD, that arguably should also be listed in this section are present elsewhere in the DSM-V.

Recent progress in genetic research on NDDs

Despite the well-known limitations of categorical approaches to the diagnosis of psychiatric syndromes, these have recently proven adequate to support dramatic advances in gene discovery in a number of conditions. Indeed, technological progress in methods to assay genetic material, combined with increasingly large patient cohorts, are now offering the first definitive, replicable insights into the genetic substrates of many of common, idiopathic NDDs. These findings are providing important avenues for in-depth studies of molecular, cellular, and circuit-level pathology. However, ironically, whereas categorical approaches to diagnosis have been sufficient to generate a run of recent seminal findings in schizophrenia, autism, epilepsy and ID, to name a few, these investigations are simultaneously undermining current diagnostic systems and blurring, if not obliterating, presumed boundaries between disorders.

The first successful systematic approaches to gene discovery in common idiopathic NDDs came as a result of the development of microarray technology that, in turn, came into widespread use as a consequence of the success of the Human Genome Project (HGP) at the turn of the millennium. This technology allowed for the parallel evaluation of initially tens of thousands of genetic markers and, later, millions of genetic markers, thus providing the first practical high-throughput assays for the unbiased detection and characterization of common sequence variation across the genome. Additionally, and equally importantly, microarray technology provided investigators with the ability to discover, and then routinely assay, the genome in search of rare and *de novo* submicroscopic variations in chromosomal structure, known as copy number variations (CNV) (for review, see Hoffman and State, 2010[10]).

One of the first successful applications of this technology to common NDDs focused on CNVs in autism[11] and

helped lead to a shift in the field from a general preoccupation with the role of common variation to the study of rare variation, particularly in simplex families (i.e. those with only a single affected individual). Moreover, the potentially disruptive nature of genomic findings for the NDD diagnostic nosology began to emerge almost immediately, with the first reports of an association of a specific CNV with autism.

Several research groups essentially simultaneously found that a structural variation on the short arm of chromosome 16 was associated with 'common, idiopathic' ASD.[12–14] This canonical 16p11.2 CNV is ~600kilobases (kb) and encompasses approximately 29 genes, 22 of which are expressed in the developing human fetal brain.[15] The CNV was observed as a *de novo* event in first study of CNVs in simplex autism.[11] Shortly thereafter, three simultaneous reports of association with the ASD phenotype were published: Kumar *et al.* described a recurrent 16p11.2 micro-deletion in four children with ASD among 712 ASD probands (0.6% frequency) whereas none were found in a cohort of 837 controls;[12];Weiss *et al.* reported five cases (0.74%) of *de novo* deletions and seven cases (0.98%) of the reciprocal duplication at this locus among 715 multiplex families, along with five deletions (0.98%) and four duplications (0.78%) in 512 children referred for evaluation of developmental delay or suspected ASD, and three deletions (1%) among 299 ASD Icelandic individuals. This was compared with the observation of two deletions (0.01%) in 18 834 unscreened Icelandic controls;[13] Marshall *et al.* reported two *de novo* deletions (0.47%), one *de novo* duplication (0.23%) and one inherited duplication (0.23%) at 16p11.2 in 427 ASD families.[14]

These reports provided independent evidence for the contribution of this CNV in autism risk; however, they were not entirely separate studies, as two of the three included overlapping subjects from the same research cohort (AGRE).[12,13] Moreover, the study from Weiss *et al.* took the (then) novel approach of considering duplications and deletions together to assess the statistical significance of their findings.[13] In addition, some of the reported pedigrees demonstrated surprising patterns of inheritance. For example, within several families with two affected siblings, only one showed the rare *de novo* 16p11.2 CNV putatively associated with ASD. In others, multiple affected family members all showed apparently *de novo* 16p11.2 mutations, suggesting germ-line mosaicism.[12,13] Finally, a case-control study of CNVs shortly thereafter failed to find additional evidence for association of 16p11.2 and ASD.[16]

However, any question about the reproducibility of these findings and the contribution of this CNV to ASD risk has been definitively answered in the ensuing 8 years. Multiple studies have now demonstrated 16p11.2 deletions in ~0.1–0.7% of cases and 16p11.2 duplications in ~0.1–0.5% of affected individuals[17,18] whereas the population base rate for both duplications and deletions has been found consistently to be at least a factor of 10 lower than this.[19,20] Moreover, Sanders *et al.* developed a novel statistical framework to evaluate the genome-wide significance of recurrent *de novo* CNV events and, in a study of 1124 ASD families from Simons Simplex Collection (SSC),[18] showed that both deletions and duplications were independently strongly associated with ASD risk at a genome-wide threshold, supporting the original findings reported by Weiss *et al.*[13].

Almost simultaneous with the demonstration that 16p11.2 deletions and duplications carried large risks in cohorts of individuals meeting research criteria for idiopathic ASD, multiple studies emerged showing the identical CNV was associated with a wide range of NDDs. For example, in a screening of 4284 patients with MR/multiple congenital anomalies, ascertained from several European centres, 22 (0.5%) had 16p11.2 deletions, with phenotypic manifestations including developmental delay (12/14), speech problems (10/14), autism (1/14), dysmorphic features (9/14) and increased body weight/obesity (5/14).[21] Similarly, Mefford *et al.* also reported six deletions (0.6%) and one duplication (0.1%) of 16p11.2 in a group of 1010 children with unexplained ID.[22] In the IQ analysis of 72 children from 16p11.2 deletion carriers and 68 intrafamilial non-carrier controls, carriers' full-scale IQs were two standard deviations (SD) lower than those of the non-carrier controls. Additionally, more than 80% of the carriers exhibited psychiatric disorders, including ASD (15% in paediatric carriers), 50% had obesity by the age of 7 years and 24% had a seizure disorder.[23]

Similarly, the finding of an association with body weight has been replicated in multiple studies: Bochyjiva *et al.* reported on 16p11.2 deletions in five patients (two *de novo* cases, accompanied by mild developmental delay) with severe early-onset obesity out of 300 Caucasian cases (1.7%) vs. two out of 7366 controls (0.03%); they evaluated an internal replication sample and identified two additional deletion cases in 1072 patients with severe obesity alone (0.2%).[24] Using multiple European cohorts ascertained for ID and/or obesity, Walter *et al.* also reported 31 cases of heterozygous deletion of 16p11.2 out of 4259 cases (0.73%).[25] In subsequent genome-wide association studies (GWAS) data for 16 053 individuals from eight European cohorts, Walter *et al.* also reported 19 similar deletions among individuals with obesity whereas none was observed in lean controls.[25] These findings were replicated in 557 cases of severe childhood obesity and in 645 control samples, along with the association between the increase of body mass index (BMI) and 16p11.2 deletion in a

general population of adults ($N = 5213$).[26] Sanders *et al.* found an inverse relationship between duplications and deletions and BMI, with increased body weight associated with deletions and decreased associated with the duplication;[18] replications of this finding were subsequently published by Jacquemont *et al.* and Steinberg *et al.*[19,27] Moreover, similar to these observations, head circumference has also associated with both the deletion and duplication in a reciprocal fashion: head circumference is larger in patients with the micro-deletion than in patients with the micro-duplication.[28]

Importantly, studies have also emerged suggesting that 16p11.2 CNVs are a risk factor for schizophrenia. Four large-scale case-control studies have identified, and reliably replicated, an approximately 10-fold increase of 16p11.2 duplications in individuals with schizophrenia, schizoaffective disorders or bipolar disorders, when compared with control groups.[19,28,31,32] This finding is even more striking in a cohort of individuals affected with childhood onset schizophrenia (COS): Ahn *et al.* reported two cases (1.6%) of 16p11.2 duplication inherited from their fathers in a total of 126 COS families; this is more frequent than in adult onset schizophrenia (0.3%).[33] Additionally, Guha *et al.* identified 13 cases of 16p11.2 deletions out of 13 850 cases with schizophrenia/schizoaffective disorder (0.094%) and 3 from 19 954 controls (0.015%), resulting in an odds ratio (OR) of 6.25.[20] These are particularly interesting observations, given the historical conceptualization of autism as childhood type schizophrenia and schizophrenic reaction, codified in both DSM I and II.[29,30] However, this was followed by a decisive swing away from this notion in later versions of the DSM which stressed key differences in symptomatology and natural history between autism and schizophrenia.

In sum, 16p11.2 CNVs serve as an early and replicated example of a single canonical locus carrying substantial risks for a wide array of phenotypes, including developmental delay, ID, schizophrenia, bipolar disorder, ASD, specific language impairment and variations in body weight and head circumference. Moreover, the finding of a base rate in carriers suggests that the presence of the CNV may be associated with no psychiatric/developmental abnormalities. Importantly, similar phenomena have been reported now for a large number of structural variations. Among these are CNVs mapping to 1q21.2, 3q29, 7q11.23, 7q36.3, 15q11.2, 15q13.3, 16p13.11, 17p12, 17q12, and 22q11.21. Indeed it has become the rule rather than the exception that CNV risks cross diagnostic boundaries.[34]

Interestingly, the same phenomenon may now be playing out with regard to NDD-associated single nucleotide mutations observed in whole exome and whole genome sequencing. Here, the contribution of *de novo* variation in NDDs is increasingly clear[35–42] and there have already been examples of mutations that were first associated with a single condition which are now being found associated with disorders that were formerly considered to be disparate, for example SCN2A in ASD and epilepsy (Table 1).[35,36,38] At present, it is difficult to discern whether the relatively modest overlap seen for point mutations, compared with the tremendous overlap in risk seen for CNVs, is a consequence of the infancy of exome and genome studies or the tremendous genetic heterogeneity of the disorders being investigated, and/or a reflection of more specific risks attending point mutations vs. CNVs that often encompass multiple genes.

The overlap among risks and the blurring of diagnostic boundaries has not been restricted to rare variations. Recent studies of common single nucleotide polymorphisms (SNPs) have shown that a risk allele for one disorder may carry comparable risks for another. The Psychiatric Genomics Consortium (PGC) examined cross-disorder effects of genome-wide significant loci previously identified for bipolar disorder and schizophrenia, using GWAS data of SNPs for 33 332 cases of ASD (4788 trios, 161 cases and 526 controls), ADHD (1947 trios, 840 cases and 688 controls), bipolar disorder (6990 cases and 4820 controls), schizophrenia (9379 cases and 7736 controls), major depressive disorders (9227 cases and 7383 controls) and 27 888 controls. The investigators reported: (i) genome-wide significance in intronic SNPs within ITIH3 and AS3MT, along with SNPs at two L-type voltage-gated calcium channel subunits, CACNA1C and CACNB2; (ii) meta-analyses for these four SNPs provided the best fit model for all five disorders in three SNPs and, for the remaining one SNP, the best fit model was limited to bipolar disorder and schizophrenia (adult onset disorders); and (iii) variations in calcium-channel activity genes seem to have pleiotropic effects on both childhood and adulthood onset disorders.[43] Subsequently, in the same datasets, the PGC used common SNPs to estimate correlations of genetic variations within and covariations between disorders. The genetic correlations were: high between schizophrenia and bipolar disorder [0.68 ± 0.04 standard error (SE)]; moderate between schizophrenia and major depressive disorder (0.43 ± 0.06 SE), bipolar disorder and major depressive disorder (0.47 ± 0.06 SE), and ADHD and major depressive disorder (0.32 ± 0.07 SE); significant but low between ASD and schizophrenia (0.16 ± 0.06 SE); and non-significant for other pairs of disorders, as well as between psychiatric disorders and controls. These results provide empirical evidence that disorders usually diagnosed after childhood (schizophrenia, bipolar disorder and major depressive disorder) have shared genetic aetiology, whereas sharing

Table 1. Recurrent non-synonymous *de novo* point mutations in ASD, schizophrenia and epilepsy[a]

Gene	Epilepsy		ASD		Schizophrenia		Study
	mut/prob	*P*-value	mut/prob	*P*-value	mut/prob	*P*-value	
GABRB3	4/264	$4.1X10^{-10}$					[35]
ALG13	2/264[§]	$7.8X10^{-12}$					[35]
SCN2A	2/264[§]	$1.1X10^{-9}$	3/575	<0.05			[35], [36], [38]
CHD8			2/122	$6.9X10^{-5}$			[37]
NTNG1			2/122	$1.2X10^{-6}$			[37]
GRIN2B			3/1703	$1.9X10^{-4}$			[37]
LAMC3			2/1703	$3.4X10^{-2}$			[37]
SCN1A			2/1703	$1.5X10^{-2}$			[37]
KATNAL2			2/375	<0.05			[38]
FMRP-associate genes			14/59 LGD in 343 prob	0.006			[39]
KIPREL3			2/40 MZ ASD twin				[40]
GPR98			2/40 MZ ASD twin				[40]
CUL3			2/233				[37], [41]
EPHB2			2/278				[36], [41]
LAM2					2/231	0.017	[42]
DPYD					2/231	0.017	[42]
TRRAP					2/231	0.017	[42]
VPS39					2/231	0.017	[42]

mut, number of mutations; prob, number of probands.

[a]Whole exome and genome studies conducted in an unbiased, hypothesis-free way were included in the review ($N = 8$ studies), and the recurrent (≥ 2), non-synonymous *de novo* point mutations were selected for the review in the table. *P*-values were computed using simulation models or likelihood model in each paper. Two whole genome studies did not provide *P*-values (studies 40, 41).

[§]Two *de novo* mutations occurred at the same position.

common variants between these disorders and childhood onset disorders (ADHD and ASD) is not likely.[44]

The continuing challenges to defining a new nosology of NDDs

In short, the accumulating data suggest, perhaps not surprisingly in retrospect,[45] that genetic risks do not conform to the boundaries defined in the current, categorical diagnostic nosology. However, the sources of this variability remain to be clarified. It is commonly assumed that gene-environment interactions, stochastic events and epigenetics must play some role in these observations. Moreover, it is likely that diagnostic substitution and ascertainment bias may contribute to some degree of observed overlap, though the consistency of findings across large samples meeting stringent diagnostic criteria for a given categorical diagnosis, the dramatic differences in natural history among NDDs such as ASD and schizophrenia, and the findings of diverse outcomes in epidemiologically-derived samples all suggest that it is unlikely that these confounds explain a substantial proportion of overlap. Finally, it is likely that there are levels of phenotypic convergence among some common NDDs that are not measured or captured in current phenotyping approaches. An in-depth discussion of these very interesting possibilities is beyond the scope of this review.

Regardless of the origin of the observations, they constitute a clear challenge to the primacy of specificity, that is the notion that in order for a particular risk factor to be important and potentially useful for illuminating pathology, it must be relevant to one and only one disorder. The blurring of traditional categorical diagnostic boundaries also raises important questions about the relationship between phenotypes and genotypes. Although historically there has been widespread interest in the notion that endophenotyping would be the key to successful gene discovery, in fact, the recent empirical evidence has largely supported the alternative, specifically that broad categorical diagnoses have been sufficient to gain a foothold in the molecular landscape of many NDDs—given sufficient sample sizes and the means to comprehensively assay the genome.

Notably, there are important exceptions to this trend with regard to gene identification. For example, even in the face of sophisticated methodologies and very large patient cohorts, there has been little progress in clarifying the specific genetic factors contributing to major depressive disorder.[46] Similarly, whereas there has been considerable progress in identifying common variations in schizophrenia,[47] there have not yet been equivalent findings with

regard to ASD, where successes in replicable gene discovery so far have been restricted to studies focused on rare variations.[48] And for the most successful examples of gene discovery in NDDs, whether for common or rare variations, the percentage of risk explained still falls well short of the anticipated contribution of genetics. These examples serve as a reminder that a great deal remains to be clarified. In some cases, such as with regard to ASD, this may simply be a result of sample cohorts that are still insufficiently powered to identify common variation; it may be that the overall genetic contribution to some NDDs has been markedly overestimated; rare variation may play a much larger role in some conditions, such as major depressive disorder (MDD), than has been anticipated; and/or additional progress in these disorders may indeed require biologically informed and dimensional endophenotyping to further clarify the relationship between molecular, cellular, circuit and behavioural phenomena.

Moreover, recent genetic findings have demonstrated a very high degree of locus heterogeneity in NDDs, for rare as well as common variations.[18,36,43,44] These observations raise important questions regarding the relationship of genetics both to diagnosis and to treatment. For example, if hundreds of genes may contribute risk for a given disorder, however defined, what are the implications for both designing and administering relevant pharmacotherapies? A full discussion of these issues is beyond the scope of this review; however, multiple recent reports suggest that a wide range of individual genetic risks correspond to a much smaller set of biological processes or points of spatial and temporal convergence.[38,40,49] These findings in turn suggest that, ultimately, treatments may be more likely to target shared aetiological substrates, rather than either individual specific mutations or broad diagnostic categories.

In the face of the ongoing conceptual challenges to diagnostics schemas, the notion of pathological specificity and the nature of the relationship of genotype to phenotype, it is unlikely that the field of psychiatry is currently in a position to come to a new and lasting clinical nosology. In fact, there is good reason to expect that the immediate future will be best served by a thoughtful combination of old, and admittedly limited, categorical diagnostic approaches, combined with new and emerging approaches to characterizing phenotypes and re-defining selected diagnoses. To some extent, these efforts have already begun.

DSM-V and the continuation of categorical diagnoses

The most recent editions of the standard descriptions of psychiatric clinical syndromes, most notably DSM and International Classification of Disease (ICD), are largely agnostic about the aetiological substrates of psychiatric conditions, including NDDs. Rather, the conventional nosology continues to focus on diagnostic criteria that are relatively easy to apply in clinical settings, while also providing the bases for national and international statistics on the morbidity and/or mortality associated with these disorders. It is also important to note that, despite the limitations of these categorical approaches, they are essential for clinical practice, as well as nascent efforts at policy development and improving public health.[50]

Furthermore, these categorical systems are not completely devoid of dimensional criteria. For example, DSM-V attempts to incorporate non-categorical and biologically informed metrics by using specifiers, subtypes, severity ratings and cross-cutting symptoms in an effort to provide clinicians with a means to better capture gradients of pathology.[9] Of course, although the DSM-V effort appears intended to augment the currently existing nosology, it remains clearly grounded in a categorical framework and restricted by its attendant limitations.

Nonetheless, despite the well-known difficulties with the approach, there are examples from other areas of medicine that suggest the possibility of a peaceful coexistence between categorical and more refined diagnostic systems. Examples include breast cancer, in which the general diagnostic indicator reflects a categorical clinical syndrome, and molecular specifiers, including estrogen/progesterone receptor status, are now routinely identified and used to guide treatment.[51] Indeed, even in the psychiatric nosology, the first strains of this type of integration are beginning to emerge. For example, a child may come to clinical attention for social difficulties, be diagnosed with ASD (serving as the basis for reimbursement and the provision of services), have genetic testing and, potentially, be diagnosed with fragile X syndrome. Although at present this molecular specifier will be most useful for family genetic counselling, the potential for a personalized medical approach, based on the molecular diagnosis, may possibly be realized with the advent of new treatments such as those seen in the studies of mGLUR5 antagonists for treating fragile X syndrome.[52]

The importance of dimensional phenotypes

Despite the current reliance on, and utility of categorical approaches, most if not all human phenotypes (both typical and pathological), including cognitive, behavioural or emotional functioning, are dimensional and continuously distributed in the population. Furthermore, anomalies in these phenotypes, for example variations in IQ, social reciprocity or executive functions, are manifest to varying degrees in many distinct current diagnostic categories, including ID, ADHD, ASD, schizophrenia and epilepsy.

The broad heterogeneity of phenotypes that result from the application of current categorical classifications has created substantial challenges in psychiatric research. As it turns out, heterogeneity often imposes significant limitations on the discovery of pathophysiological mechanisms because there are often problems in determining the comparability of subjects in particular clinical studies. This leads to disparate study findings and failures to replicate what may actually be meaningful initial findings. That is not to say that categorical diagnostic systems are useless. In fact, when rigorously applied, categorical diagnostic criteria for psychiatric disorders may be advantageous in some aetiological studies by enriching the extremes of a particular phenotype. Unfortunately, this can work both ways and also increase phenotype heterogeneity, based on subject selection.

At present, the best organized and most determined approach to implementing a dimensional system for neuropsychiatric conditions is embodied in the NIMH Research Domain Criteria (RDoC) initiative.[53] This is an experimental approach that is agnostic to previous diagnostic categorical classifications, indeed cutting across traditional clinical diagnostic boundaries in an effort to reduce heterogeneity within yet-to-be-determined syndromes.[54] RDoC is a framework for research that aims to create a stronger foundation for research on pathophysiology, especially for genomics and neuroscience, which it is hoped will ultimately inform future clinical classification schemes and eventually help identify new targets for treatment development.[55]

Based on five major 'domains' and eight levels of analysis, the RDoC matrix is conceived as a model for incorporating biological and environmental data into a nosological framework (Table 2).[56] RDoC specifically allows for genetic, molecular and cellular level data to be included in a descriptive system. However, it remains unclear as to whether the 'domains' are truly relevant to clinical phenomena and how this system can be applied to describing and treating specific clinical conditions.

Research using the new DSM-V as well as the RDoC systems is in its infancy. It remains to be seen how well these two different approaches will facilitate discovery of aetiological substrates and pathophysiological mechanisms of psychiatric disorders. It is likely that neither will be sufficiently utilitarian for all purposes and that either they will need to be combined, or a new, more integrated system will need to be developed. In any case, at present there is still a need for a single dimensional nosology that can account for heterogeneity and allow for the inclusion of behavioural phenotypes as well as biological and environmental data. This is necessary in order to define specific phenotypes that will advance our understanding of the aetiologies of NDDs and other psychiatric disorders as well as lead to appropriate treatment.

Additional steps toward a new synthesis—combining dimensional diagnosis with epidemiologically-ascertained population-based samples

For a variety of reasons, most descriptive psychiatric research has focused on clinical populations. Cost, efficiency, time and other factors have led investigators to ascertain study subjects from their clinics. However, such convenience has come at a price: sampling errors are inherent to the exclusive use of clinical samples. This problem is largely unrecognized in the study of psychiatric disorders and, in particular, of NDDs.

Certainly, clinical samples can serve as a useful initial step in the process of identifying a clinical syndrome or phenotype. However, by their very nature, clinical samples are not likely to represent the entire spectrum or dimensionality of any particular syndrome. Furthermore, clinical samples are known to be biased with respect to case severity, comorbidity and factors associated with health care system access.[57,58] As a result, clinical samples provide a biased view of a clinical condition that can only be corrected with other methods of sampling. Epidemiologically ascertained, population-based samples minimize or even

Table 2. RDoC matrix (domains and analysis system)

	Units of analysis							
Domains	Genes	Molecules	Cells	Circuits	Physiology	Behaviour	Self-report	Paradigms
Negative valence systems								
Positive valence systems								
Cognitive systems								
Social processing systems								
Arousal & regulatory systems								

eliminate such biases. By including a representative sample of all individuals in a given community, epidemiological sampling allows for appreciation of the variation inherent in a clinical condition, as well as a perspective on the full dimensionality of clinical presentation and disability.

For example, in a recent comprehensive, population-based ASD prevalence study, led by one of us (Y.S.K.) using a total population approach for ascertainment, we found that ASD is much more prevalent than previously thought: 2.6%. A somewhat surprising two-thirds of the children ascertained (1.89% prevalence) with ASD were previously unidentified in the community and would not have been ascertained in a clinical sample. Moreover, the phenotypic characteristics of those with ASD found in non-clinical settings were distinct from clinically derived individuals. They had a mean IQ of 98 and less male predominance (2.5:1).[59] This study suggests that the full spectrum of the ASD phenotype is consistent with current research that demonstrates a continuous distribution of the dimensional phenotype of 'autism traits'.[60]

These observations serve to point out the importance of systematically ascertained, population-based ASD samples for ASD research because they alone permit examination of underlying ASD pathophysiology across the entire ASD spectrum, while also allowing for proper assessment of the full clinical spectrum leading to accurate behavioural phenotyping.[61–65] Moreover, as with ASD, it is likely that epidemiological samples for other NDDs will be essential in order to fully appreciate the complexity and distribution of the clinical and biological characteristics of these disorders.

Conclusion

The limitations of categorical diagnostic schema, particularly for common NDDs, are widely recognized, and there has been speculation that these approaches have dramatically hindered the search for the causes and treatments of psychiatric conditions. However, the data also speak clearly, at least with regard to genetics: the combination of advanced techniques, adequate samples and an admittedly imperfect categorical approach to diagnosis is in fact adequate to fuel key discoveries in NDDs. The conundrum as noted is that whereas recent work exploring common, rare and *de novo* variants has led to the identification of specific genetic risks for several paradigmatic disorders, at the same time these findings have called into question the boundaries established between these conditions and challenged the notion that a unique disorder (if appropriately defined and isolated) has a unique aetiology.

Although some may see this as revealing the fundamental flaws in current nosological systems, it seems more useful to conceptualize this as an ongoing and iterative process: with categorical diagnosis sufficient to provide an entry into molecular mechanisms, which then serve to help develop and refine a new, more biologically relevant nosology that allows for both genetic and phenotypic heterogeneity. The beginning strains of such a system can be found in the co-existence of RDoC and DSM-V, especially when this composite can now increasingly include the identification of specific genetic risks for NDDs. As noted above, this may portend an emerging approach not unlike that for some cancers, where a clinical/organ-based diagnosis may be used to categorize an initial presentation but the patient is then further conceptualized and, in an increasing number of cases, treated based on a combination of the clinical presentation, specific genetic markers and dimensional biological measures to guide therapy.

Funding

This work was supported by grants ES021462 (Y.S.K.) and MH089956, MH081754 and MH100027(M.W.S.).

Author contributions

Y.S.K. has checked the references for accuracy and completeness. M.S. acts as guarantor for the paper. Both authors confirm that this material has not been published previously in a substantively similar form.

Conflict of interest: None declared.

References

1. Yeargin-Allsopp M, Boyle C. Overview: the epidemiology of neurodevelopmental disorders. *Ment Retard Dev Disabil Res Rev* 2002;**8**:113–16.
2. Accardo PJ, Capute AJ. *Capute & Accardo's Neurodevelopmental Disabilities in Infancy and Childhood: Neurodevelopmental diagnosis and treatment*. Baltimore, MD: Paul H Brookes, 2008.
3. American Psychiatric Association. *DSM-IV*. Arlington, VA: American Psychiatric Association, 1994.
4. McKusick-Nathans Institute of Genetic Medicine. *Online Mendelian Inheritance in Man*. Baltimore, MD: Johns Hopkins University School of Medicine, 2013.
5. Moss J, Howlin P. Autism spectrum disorders in genetic syndromes: implications for diagnosis, intervention and understanding the wider autism spectrum disorder population. *J Intellect Disabil Res* 2009;**53**:852–73.
6. Morrow EM, Yoo SY, Flavell SW *et al*. Identifying autism loci and genes by tracing recent shared ancestry. *Science* 2008;**321**:218–23.
7. Yu TW, Chahrour MH, Coulter ME *et al*. Using whole-exome sequencing to identify inherited causes of autism. *Neuron* 2013;**77**:259–73.

8. Novarino G, El-Fishawy P, Kayserili H *et al*. Mutations in BCKD-kinase lead to a potentially treatable form of autism with epilepsy. *Science* 2012;**338**:394–97.
9. American Psychiatric Association. *DSM-V*. Arlington, VA: American Psychiatric Association, 2013.
10. Hoffman EJ, State MW. Progress in cytogenetics: implications for child psychopathology. *J Am Acad Child Adolesc Psychiatry* 2010;**49**:736–51; quiz 856–57.
11. Sebat J, Lakshmi B, Malhotra D *et al*. Strong association of de novo copy number mutations with autism. *Science* 2007;**316**:445–49.
12. Kumar RA, KaraMohamed S, Sudi J *et al*. Recurrent 16p11.2 microdeletions in autism. *Hum Mol Genet* 2008;**17**: 628–38.
13. Weiss LA, Shen Y, Korn JM *et al*. Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med* 2008;**358**:667–75.
14. Marshall CR, Noor A, Vincent JB *et al*. Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet* 2008;**82**:477–88.
15. Simons Variation in Individuals Project (Simons VIP). A genetics-first approach to studying autism spectrum and related neurodevelopmental disorders. *Neuron* 2012;**73**:1063–67.
16. Glessner JT, Wang K, Cai G *et al*. Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* 2009;**459**:569–73.
17. Shen Y, Dies KA, Holm IA *et al*. Clinical genetic testing for patients with autism spectrum disorders. *Pediatrics* 2010;**125**:e727–35.
18. Sanders SJ, Ercan-Sencicek AG, Hus V *et al*. Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron* 2011;**70**:863–85.
19. Jacquemont S, Reymond A, Zufferey F *et al*. Mirror extreme BMI phenotypes associated with gene dosage at the chromosome 16p11.2 locus. *Nature* 2011;**478**:97–102.
20. Guha S, Rees E, Darvasi A *et al*. Implication of a rare deletion at distal 16p11.2 in schizophrenia. *JAMA Psychiatry* 2013;**70**: 253–60.
21. Bijlsma EK, Gijsbers AC, Schuurs-Hoeijmakers JH *et al*. Extending the phenotype of recurrent rearrangements of 16p11.2: deletions in mentally retarded patients without autism and in normal individuals. *Eur J Med Genet* 2009;**52**:77–87.
22. Mefford HC, Cooper GM, Zerr T *et al*. A method for rapid, targeted CNV genotyping identifies rare variants associated with neurocognitive disease. *Genome Res* 2009;**19**:1579–85.
23. Zufferey F, Sherr EH, Beckmann ND *et al*. A 600 kb deletion syndrome at 16p11.2 leads to energy imbalance and neuropsychiatric disorders. *J Med Genet* 2012;**49**:660–68.
24. Bochukova EG, Huang N, Keogh J *et al*. Large, rare chromosomal deletions associated with severe early-onset obesity. *Nature* 2010;**463**:666–70.
25. Walters RG, Jacquemont S, Valsesia A *et al*. A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. *Nature* 2010;**463**:671–75.
26. Walters RG, Coin LJ, Ruokonen A *et al*. Rare genomic structural variants in complex disease: lessons from the replication of associations with obesity. *PLoS One* 2013;**8**:e58048.
27. Steinberg S, de Jong S, Mattheisen M *et al*. Common variant at 16p11.2 conferring risk of psychosis. *Mol Psychiatry* 2014;**19**: 108–14.
28. McCarthy SE, Makarov V, Kirov G *et al*. Microduplications of 16p11.2 are associated with schizophrenia. *Nat Genet* 2009;**41**: 1223–27.
29. American Psychiatric Association. *DSM-I*. Arlington. VA: American Psychiatric Association, 1952.
30. American Psychiatric Association. *DSM-II*. Arlington, VA: American Psychiatric Association, 1968.
31. Levinson DF, Duan J, Oh S *et al*. Copy number variants in schizophrenia: confirmation of five previous findings and new evidence for 3q29 microdeletions and VIPR2 duplications. *Am J Psychiatry* 2011;**168**:302–16.
32. Grozeva D, Conrad DF, Barnes CP *et al*. Independent estimation of the frequency of rare CNVs in the UK population confirms their role in schizophrenia. *Schizophr Res* 2012;**135**:1–7.
33. Ahn K, Gotay N, Andersen TM *et al*. High rate of disease-related copy number variations in childhood onset schizophrenia. *Mol Psychiatry* 2013 May 21. doi: 10.1038/mp.2013.59.
34. Malhotra D, Sebat J. CNVs: harbingers of a rare variant revolution in psychiatric genetics. *Cell* 2012;**148**:1223–41.
35. Allen AS, Berkovic SF, Cossette P *et al*. De novo mutations in epileptic encephalopathies. *Nature* 2013;**501**:217–21.
36. Sanders SJ, Murtha MT, Gupta AR *et al*. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* 2012;**485**:237–41.
37. O'Roak BJ, Vives L, Girirajan S *et al*. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 2012;**485**:246–50.
38. Neale BM, Kou Y, Liu L *et al*. Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* 2012;**485**: 242–45.
39. Iossifov I, Ronemus M, Levy D *et al*. De novo gene disruptions in children on the autistic spectrum. *Neuron* 2012;**74**:285–99.
40. Michaelson JJ, Shi Y, Gujral M *et al*. Whole-genome sequencing in autism identifies hot spots for de novo germline mutation. *Cell* 2012;**151**:1431–42.
41. Kong A, Frigge ML, Masson G *et al*. Rate of de novo mutations and the importance of father's age to disease risk. *Nature* 2012;**48**:471–75.
42. Xu B, Ionita-Laza I, Roos JL *et al*. De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. *Nat Genet* 2012;**44**:1365–69.
43. Smoller JW, Craddock N, Kendler K *et al*. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* 2013;**381**:1371–79.
44. Lee SH, Ripke S, Neale BM *et al*. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* 2013;**45**:984–94.
45. State MW, Levitt P. The conundrums of understanding genetic risks for autism spectrum disorders. *Nat Neurosci* 2011;**14**: 1499–506.
46. Shyn SI, Shi J, Kraft JB *et al*. Novel loci for major depression identified by genome-wide association study of sequenced treatment alternatives to relieve depression and meta-analysis of three studies. *Mol Psychiatry* 2011;**16**:202–15.
47. Mowry BJ, Gratten J. The emerging spectrum of allelic variation in schizophrenia: current evidence and strategies for the identification and functional characterization of common and rare variants. *Mol Psychiatry* 2013;**18**:38–52.

48. Devlin B, Melhem N, Roeder K. Do common variants play a role in risk for autism? Evidence and theoretical musings. *Brain Res* 2011;**1380**:78–84.
49. Willsey AJ, Sanders SJ, Li M *et al*. Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell* 2013;**1**:997–1007.
50. Hyman SE. Can neuroscience be integrated into the DSM-V? *Nat Rev Neurosci* 2007;**8**:725–32.
51. NCI. Tumor Markers. 2013. NCI FactSheet, 2013. Available from http://www.cancer.gov/cancertopics/factsheet/detection/tumor-markers (03 March 2014, date last accessed).
52. Krueger DD, Bear MF. Toward fulfilling the promise of molecular medicine in fragile X syndrome. *Ann Rev Med* 2011;**62**:411–29.
53. NIMH. RDoC. 2013. Research Priorities 2013. Available from: http://www.nimh.nih.gov/research-priorities/rdoc/index.shtml (03 March 2014, date last accessed).
54. Carpenter WT. RDoC NIMH, RDoC and DSM-V: What's the fuss? *Schizophr Bull* 2013;**39**:945–46.
55. Insel T, Cuthbert B, Garvey M *et al*. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am J Psychiatry* 2010;**167**:748–51.
56. NIMH. NIMH RDoC Matrix. 2013. Rearch Priorities RDoC 2013. Available from http://www.nimh.nih.gov/research-priorities/rdoc/nimh-research-domain-criteria-rdoc.shtml#toc_matrix (03 March 2014, date last accessed).
57. Di Sclafani V, Finn P, Fein G. Treatment-naive active alcoholics have greater psychiatric comorbidity than normal controls but less than treated abstinent alcoholics. *Drug Alcohol Depend* 2008;**98**:115–22.
58. Berkson J. Limitations of the application of the fourfold table analysis to hospital data. *Biometrics* 1946;**2**:47–53.
59. Kim YS, Leventhal BL, Koh YJ *et al*. Prevalence of autism spectrum disorders in a total population sample. *Am J Psychiatry* 2011;**168**:904–12.
60. Constantino JN, Todd RD. Intergenerational transmission of subthreshold autistic traits in the general population. *Biol Psychiatry* 2005;**57**:655–60.
61. Rutter M, Andersen-Wood L, Beckett C *et al*. Quasi-autistic patterns following severe early global privation. English and Romanian Adoptees (ERA) Study Team. *J Child Psychol Psychiatry* 1999;**40**:537–49.
62. Manolio TA. Cohort studies and the genetics of complex disease. *Nat Genet* 2009;**41**:5–6.
63. Aulchenko YS, Ripatti S, Lindqvist I *et al*. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nat Genet* 2009;**41**:47–55.
64. Kathiresan S, Willer CJ, Peloso GM *et al*. Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet* 2009–65.
65. Sabatti C, Service SK, Hartikainen AL *et al*. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat Genet* 2009;**41**:35–46.