

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:<https://orca.cardiff.ac.uk/id/eprint/94486/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

O'Donovan, Michael Conlon and Owen, Michael John 2016. The implications of the shared genetics of psychiatric disorders. *Nature Medicine* 22 , pp. 1214-1219. 10.1038/nm.4196

Publishers page: <http://dx.doi.org/10.1038/nm.4196>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



1
2 **Michael C O'Donovan* and Michael J Owen**

3
4
5 **The implications of the shared genetics of psychiatric disorders**

6
7
8
9
10 MRC Centre for Neuropsychiatric Genetics and Genomics,
11 Division of Psychiatry and Clinical Neurosciences
12 Cardiff University School of Medicine
13 Cardiff, UK
14 CF24 4HQ
15 *corresponding author; email: odonovanmc@cardiff.ac.uk

16
17
18 **Abstract**

19
20 Recent genomic studies have revealed the highly polygenic nature of psychiatric disorders
21 including schizophrenia, bipolar disorder, and major depressive disorder. Many of the
22 individual genetic associations are shared across multiple disorders in a way that points to
23 extensive biological pleiotropy, and further challenges the biological validity of existing
24 diagnostic approaches. Here, we argue it is unlikely that risk alleles exist that are specific to
25 a single diagnostic category. We also highlight some of the important clinical repercussions
26 of pleiotropy.

27
28
29
30 **Introduction**

31
32 Psychiatric disorders represent 21st Century Medicine's greatest global challenge. They are
33 the major cause, worldwide, of non-fatal burden of disease¹. They account for around 30%
34 of all years lived with disability, a contribution that is rising, especially in developing countries
35 as the burden moves from communicable to non-communicable diseases¹. With a life time
36 prevalence greater than 10%, major depressive disorder accounts for a higher burden of
37 disability than any other disorder, while schizophrenia, which contributes less to global
38 burden due to its lower prevalence (around 1%), is the most severely disabling of all medical
39 conditions². It is stating the obvious that we need to develop and deliver more effective
40 psychiatric prevention and treatment, but despite years of effort there have been few

41 significant advances. There are a number of reasons for this, but prominent among them has
42 been our lack of understanding of aetiology and pathogenesis, compounded by our reliance
43 on observational and syndromic systems of diagnosis and classification.

44
45 Concerns about psychiatric diagnosis and classification have been thrown into sharp relief
46 by recent genomic studies that appear to show that risk alleles tend not to be specific to any
47 particular disorder. In this perspective, we discuss the nature and extent of the evidence for
48 shared risk alleles across psychiatric disorders and interpret that evidence within the context
49 of how psychiatric diagnoses are made. We consider whether it is likely that risk variants for
50 specific disorders exist, and how future studies might usefully illuminate alternative
51 genotype-phenotype relationships. We also consider some of the clinical implications
52 emerging from pleiotropy. The focus of our discussion is the major psychiatric disorders such
53 as schizophrenia, major mood disorders, autism spectrum disorders (ASD) and attention
54 deficit hyperactivity disorder (ADHD), disorders for which research has faced conceptual and
55 practical challenges that place them apart from dementias and from acute disturbances of
56 mental function that are secondary to trauma, toxicity, or medical conditions.

57

58 **Psychiatric Diagnosis**

59

60 *The Making of Psychiatric Diagnosis*

61 Psychiatric diagnoses are made on the basis of patient descriptions of their subjective
62 experiences (e.g. energy, mood, perception, beliefs, appetite), and from observations of
63 behaviour (e.g. bizarre activity, attention, self-care, social interaction) made by clinicians or
64 reported by informants (e.g. family or carers, neighbours, teachers). Other factors are taken
65 into consideration including functional impairment, developmental trajectory, and outcome.
66 Ultimately a diagnosis is assigned to individuals who exhibit a minimum number of
67 symptoms, behaviours, or outcomes, usually for a minimum period of time, with the proviso
68 they do not meet criteria that exclude that diagnosis. The exclusion criteria are often
69 subjective, requiring clinicians to judge that the clinical picture is not 'better accounted for' by
70 another diagnosis, or that the picture is not 'clearly caused by' the effects of a psychoactive
71 agent for example. In clinical practice, experience and intuition play a role, although semi-
72 standardized data acquisition tools and operationalized diagnostic criteria have been
73 developed to minimize the impact of these subjective factors. These are used primarily in
74 research, but they are sometimes employed to aid diagnosis in the clinic (e.g. the Autism
75 Diagnostic Observation Schedule).

76

77 *The Validity of Psychiatric Diagnosis*

78 As research data have accumulated it has become clear that the boundaries between
79 diagnostic groups and between illness and wellness are not clear-cut, there is considerable
80 heterogeneity within diagnostic categories, patients often have the clinical features of more
81 than one disorder³, and the preponderance of those features in a particular individual can
82 change markedly over time and development. Even with the most fastidious application of
83 diagnostic criteria, there is no avoiding the fact that none of the clinical features are
84 pathognomonic. For example, the occurrence of psychotic symptoms such as hallucinations
85 and delusions, mood changes, and alterations in speech, activity level, behaviour and sleep,
86 can indicate either a diagnosis of schizophrenia or bipolar disorder (BD). The frequent co-
87 occurrence of symptoms that could imply either major diagnostic label has led to a third
88 category, schizoaffective disorder. Archetypal versions of each diagnosis exist, but for a
89 large number of people, the distinction is based on relatively subjective judgments about the
90 duration, quality, and severity of component signs and symptoms⁴. Outcome within
91 diagnostic groups also varies widely, for example some people with a diagnosis of
92 schizophrenia remain chronically symptomatic and impaired while others make a complete
93 recovery⁵. Finally, as better-powered epidemiological studies have been carried out, it has
94 become clear that the relatives of an individual with one psychiatric diagnosis are at
95 increased risk for other diagnoses, undermining the genetic validity of current diagnostic
96 approaches⁶.

97

98 *Biomarkers*

99 There have been extensive efforts to identify biomarkers indexing pathogenic mechanisms,
100 including studies of blood markers (e.g. metabolites, cytokines, cortisol suppression),
101 behavioural and cognitive measures, and various neuroimaging modalities⁷. However, this
102 work has failed to deliver markers that can reliably distinguish between diagnoses, or to
103 identify disease subgroups and, currently, there are no biomarkers in routine clinical use.
104 For example, despite the extensive use of ever more sophisticated neuroimaging
105 approaches, no measures have emerged that can separate people with a particular
106 diagnosis from healthy individuals, much less distinguish between those with different
107 diagnoses⁸.

108

109

110 **Molecular Genetic Findings in Psychiatry**

111

112 The robust identification of risk factors for psychiatric disease as indicated by DNA variation
113 has been eagerly awaited for the insights this might provide into the basic biological
114 architecture of, and relationships between, psychiatric phenotypes, as well as for its

115 contributions to understanding disease mechanisms. In the last few years, genomic studies
116 have begun to identify risk alleles in large numbers; although success has largely been
117 confined to ASD, schizophrenia, and, to a lesser extent, BD and Major Depressive Disorder
118 (MDD). Other psychiatric phenotypes have yet to be subjected to large-scale genome-wide
119 studies.

120

121

122 In ASD, the evidence implicating specific risk genes comes primarily from mutations that
123 occur *de novo* in the form of large insertion-deletion mutations called copy number variants
124 (CNVs) or rare coding variants (RCVs) that change the DNA sequence at a single, or a few,
125 nucleotides. A recent synthesis of the ASD data (5,563 cases for *de novo* RCVs, 4687 cases
126 for CNVs) reported high confidence associations to 65 genes and an additional 6 CNV loci⁹.
127 All loci identified thus far confer large effects on risk, but with population frequencies less
128 than one in a thousand, however this might simply reflect low power to detect smaller effect
129 sizes. It should also be noted that there is emerging evidence that common genetic variation
130 makes a substantial contribution to the variance in liability to ASD¹⁰ although, individual
131 common alleles have not yet been robustly implicated.

132

133 In schizophrenia, identified risk alleles span the full spectrum of frequencies. The largest
134 analysis of genome-wide association (GWAS) data (up to 36,989 cases and 113,075
135 controls) identified a total of 108 loci containing common alleles while that of rare CNVs
136 (12,029-21,269 cases; 24,815- 81,821 controls) identified 11 strongly supported loci¹². The
137 latter was largely based on a meta-analysis of candidate CNVs, and a systematic genome
138 wide CNV meta-analysis is awaited. Exome sequencing studies in schizophrenia have been
139 smaller than those in ASD, and the evidence for RCVs is largely restricted to enrichments in
140 pathways rather than specific genes^{13,14}, although recently, a meta-analysis (4,264
141 schizophrenia cases, 9,343 controls, 1,077 parent-proband trios) obtained genome-wide
142 significant association between schizophrenia and Loss-of-Function (LoF) RCVs in a gene
143 which encodes the histone methyltransferase SETD1A¹⁵. That study also reported a specific
144 mutation in *SETD1A* that occurred in people with the disorder as a *de novo* mutation at a
145 frequency far in excess of that expected by chance, providing confidence for pathogenicity of
146 that specific mutation.

147

148 In BD, GWAS and rare variant datasets are smaller than those of schizophrenia. The largest
149 GWAS study (9,747 patients and 14,278 controls) identified 5 risk loci while, at the rare
150 variant end of the spectrum, the only finding that meets a statistical threshold equivalent to
151 genome wide significance is a duplication CNV at 16p.11.2¹⁶. Finally, a recent GWAS¹⁷

152 based upon MDD, as self-reported by customers of a consumer genetics company, identified
153 15 loci for the disorder. It is particularly notable that only 15 loci were identified in a study
154 including up to 130,620 cases and 347,620 controls. This underscores the fact that, while
155 sample size may be critical for discovery genetics, it is not the only factor. Differences in
156 other properties of disorders (e.g. disease prevalence, heterogeneity, phenotype definition,
157 variance in risk contributed by individual alleles) can have a major impact.

158

159

160 **Pleiotropy**

161

162 *The Nature of Pleiotropy*

163 The meaning of pleiotropy (Figure 1) depends on context^{18,19}. We refer to genic pleiotropy
164 when altered function of a gene influences multiple traits (note the term trait includes
165 phenotypes that are not necessarily abnormal or disorders). Allelic pleiotropy, a subtype of
166 genic pleiotropy, occurs when the same gene variant influences multiple traits. This is
167 exemplified by phenylketonuria (PKU) in which causative mutations are pleiotropic for
168 intellectual disability, lack of pigmentation, as well as various metabolic changes that can be
169 measured in the blood. These two forms of biological pleiotropy, genic and allelic, suggest
170 shared biology between disorders, but this is not the only explanation.

171

172 Mediated pleiotropy occurs when an allele influences two traits, but its effects on one are
173 secondary to more direct effects on the other. For example genetic variation at the *fat mass*
174 *and obesity associated (FTO)* locus is pleiotropic for body mass index (BMI) and type 2
175 diabetes (T2D), but the effects on T2D are secondary to those on BMI. In the case of PKU,
176 the effects on intellectual function and pigmentation are mediated by the effects on the
177 metabolic traits. As in these examples, mediated pleiotropy can be informative for
178 understanding causal pathways to disease and, as we shall see, is often implicitly assumed
179 in endophenotype studies, but the mediating relationship between the two traits can be
180 complex and it does not necessarily imply that the two phenotypes share biological
181 mechanisms.

182

183 There are also numerous sources of false or pseudo pleiotropy. Pseudo pleiotropy can arise
184 as a result of imprecision in gene mapping where two phenotypes are influenced by different
185 genes in close proximity (Figure 1) but it can also arise from poor study design, associations
186 that are due to chance (type II errors), and publication biases favouring reports of overlaps.

187

188 *Pleiotropy in Psychiatry and Developmental Disorders*

189 Evidence for cross disorder effects of genetic variation has come from studies showing that
190 CNVs that influence risk for schizophrenia also often do so for ASD, intellectual disability (ID)
191 developmental delay (Figure 2), and ADHD²⁰. The majority of these apparently pleiotropic
192 CNVs are multigenic, and therefore we cannot exclude pseudo pleiotropy in which distinct
193 genes within the CNV cause each associated phenotype (Figure 1). However, the
194 observation that every CNV known to increase risk of schizophrenia also does so for ID²¹
195 makes co-localization alone an unlikely explanation. Moreover, the only 'single gene' CNV
196 that is unequivocally associated with schizophrenia, deletion of the gene *NRXN1* encoding
197 the pre-synaptic protein neurexin 1 is also associated with ASD and with ID²². Sequencing
198 studies have shown that as a group, genes impacted by LoF *de novo* mutations in
199 schizophrenia are enriched for those affected by this same class of mutation in people with
200 ASD and ID¹³. Moreover, several genes have been definitively implicated by *de novo* LoF
201 mutations in each of developmental delay and ASD^{9,23}, while at an even finer level of
202 resolution, the same LoF mutation in *SETD1A* that contributes high risk to schizophrenia
203 also does so for severe ID and developmental delay¹⁵.

204
205 The hypothesis of true pleiotropy in psychiatric and developmental disorders is also
206 supported by common variants identified by GWAS. The International Schizophrenia
207 Consortium (ISC) showed that hundreds, and perhaps thousands, of common alleles that
208 increase risk for schizophrenia also do so for BD²⁴ and it is now clear they also do so for
209 MDD, and to a lesser extent, ASD, ADHD, Anorexia Nervosa, Obsessive Compulsive
210 Disorder (Figure 3), as well as personality traits such as neuroticism²⁵⁻²⁷. A problem with
211 inferring biological pleiotropy from GWAS is that the functional alleles (i.e. the alleles that
212 changes function or expression of the gene and directly cause the association) responsible
213 for the vast majority of the GWAS associations have not been identified. It is therefore
214 possible that for any single cross-disorder association, different functional variants within the
215 same or different genes might be responsible. However, the substantial genetic correlations
216 between pairs of psychiatric phenotypes (Figure 3) are less readily explained by pseudo
217 pleiotropy as this would require different functional alleles to be systematically and
218 consistently tagged by the same GWAS allele across large numbers of loci.

219
220 Taking the genomic data as a whole, true pleiotropy is by far the most parsimonious
221 explanation for the majority of published cross disorder effects, and most of the findings
222 support extensive allelic pleiotropy. A proviso here is that we must exclude mediated
223 pleiotropy as an explanation. By definition, for one trait to be secondary to (or mediated by)
224 another, the mediating trait must occur first. It follows that childhood onset disorders (e.g.
225 ADHD) cannot be mediated by disorders with typically later ages of onset (e.g.

226 schizophrenia, MDD). However, it is theoretically possible that the converse is true, and that
227 where alleles are pleiotropic for ID, schizophrenia, and ASD, ID is the primary phenotype
228 influenced by those alleles, and that having ID causally increases risk of ASD and
229 schizophrenia. There is certainly evidence that CNVs and *de novo* LoF mutations occur
230 more frequently in people with psychiatric disorders who additionally have cognitive
231 impairment^{9,13,28}, an observation that has sometimes been interpreted as indicative of
232 pleiotropy mediated through ID. However, this pattern of co-morbidity is not sufficient to
233 establish mediated pleiotropy, indeed it is to be expected in cases where mutations have
234 direct effects on two phenotypes. There are also powerful arguments against mediated
235 pleiotropy as the sole explanation for this. First, in ASD, LoF *de novo* mutations tend to
236 occur in the same sets of genes in probands with and without intellectual disability⁹. Second,
237 at *SETD1A*, although LoF mutations are associated with both ID and schizophrenia, ID is not
238 a prerequisite for schizophrenia in mutation carriers¹⁵. Third, ID is not universally seen in
239 people with schizophrenia who carry *de novo* CNVs that are pleiotropic for both disorders²⁸.
240 Fourth, in the only study we are aware of that has explicitly undertaken a formal mediation
241 analyses based on a rare variant, the 22q11 deletion CNV was found to have independent
242 effects on cognitive and psychiatric traits (e.g. ADHD and ASD)³⁰. The rare variant data are
243 therefore inconsistent with the hypothesis that cross disorder findings are explained by
244 mediated rather than allelic pleiotropy. The common variant findings are more complex, and
245 will be considered further below.

246

247 ***Pleiotropy in the context of complex disorders***

248

249 Pleiotropy is a challenging phenomenon in the context of highly polygenic disorders.
250 Consider CNVs associated with at least two clinical outcomes, schizophrenia and intellectual
251 disability, as well as being present in apparently unaffected carriers with no clinical
252 phenotype. It has recently been shown³¹ that clinically unaffected CNV carriers perform
253 worse on a range of measures of cognitive performance than do non-carrier controls, but
254 better than people with either of the clinical diagnoses associated with the CNVs. Cognitive
255 phenotyping therefore empirically demonstrates that CNVs impact on liability to quantitative
256 traits that are overlooked when the only definition of 'affected' is that of a clinical diagnosis.
257 What determines the final manifestations of increased liability in CNV carriers is not well
258 understood, but an individual's burden of common schizophrenia risk alleles is one important
259 factor³². What might then be perceived as pleiotropic manifestations of a particular mutation
260 (e.g. a CNV) may in fact more generally represent the net effects of an individual's polygenic
261 and environmental background on multiple traits representing various domains of brain
262 function.

263

264 **Specific genes for psychiatric diagnoses**

265

266 Whether it is possible to link genotype to psychiatric phenotype is generally couched in
267 terms of linear relationships between a gene and a single categorical diagnosis. In our
268 opinion, the evidence summarized above suggests the outlook for relating genotype and
269 phenotype in this way is not promising, although we recognize that there is a bias towards
270 observing pleiotropy since studies are better powered to identify genetic similarities rather
271 than differences.

272

273 We do not suggest that risk alleles impact on psychiatric outcomes indiscriminately. For
274 example duplication at 22q11 increases risk of ID and ASD, but is neutral for bipolar
275 disorder, and protective for schizophrenia³³. Damaging rare mutations play a greater role in
276 ID than in schizophrenia, in schizophrenia compared with mood disorder, and in psychiatric
277 disorders with comorbid cognitive impairment¹⁶. With regard to common alleles, although
278 many psychiatric disorders are genetically correlated, the degree of correlation between
279 diagnostic classes is usually less than the degree of within disorder correlation^{26,34}. These
280 observations suggest that current diagnostic schemes do to some extent capture groups
281 whose members have more in common with each other than they do with members of a
282 general class 'psychiatric disorder'. However, until we can directly measure liability, it is
283 impossible to distinguish the phenotypic heterogeneity arising from true pleiotropic effects of
284 a specific allele (even an allele of large effect) from that resulting from a person's unique
285 blend of risk factors. Directly measuring liability remains a distant goal; for now, identifying
286 alternative approaches to patient stratification that index liability better than current
287 diagnostic categories, and therefore might link more specifically to particular genotypes, is a
288 more realistic aim. Some approaches to doing so are outlined in Box 1.

289

290

291 **Implications of Pleiotropy.**

292

293 The current system of psychiatric classification is not optimal, and alternative approaches
294 are urgently required for clinical and fundamental research. The genetic findings do not,
295 however, imply a similar urgency for fundamental changes in clinical practice as they do not
296 provide the basis for a system with clear clinical value. Given the complexity of the
297 relationships between disorders, and the likelihood that people with psychiatric illnesses
298 differ quantitatively on multiple dimensions of function rather than categorically, seeking hard
299 categorical boundaries that validly reflect aetiology seems a fool's errand. Ultimately, we

300 suspect the advances in genomic research will allow mapping between pathophysiological
301 processes and domains of brain function (perhaps those outlined in RDoC, perhaps not) and
302 between domains of brain function and the clinical picture and in doing so, will allow clinical
303 measurements (for example types of cognitive test, brain imaging) that highlight
304 perturbations that are pertinent to, and suggest interventions for, particular groups of
305 patients. But what measures are likely to best achieve this, much less how to implement
306 them in a clinical setting, is far from clear. Nevertheless, even now, the pervasive nature of
307 shared risk factors, pleiotropy, and arbitrary diagnostic boundaries between disorders has
308 clinical implications.

309
310 As clinicians, we recognize the utility of diagnostic boundaries for therapeutic decision
311 making, communication, and predicting (in a general way) certain outcomes and we do not
312 suggest that clinicians abandon diagnosis using existing categories. However, rigid
313 adherence to categories makes it easy to either overlook co-morbidity or, where it is
314 detected, to inappropriately ascribe it to a diagnosis that has greater weight in the current
315 diagnostic hierarchy. As a result, co-morbid syndromes may not be optimally treated. Given
316 that pleiotropy implies that a person with one syndrome is at enhanced risk for a second
317 syndrome, far from implying lax assessment, pleiotropy emphasizes the need for detailed
318 on-going clinical monitoring, and assessments that go beyond the bare requirements of
319 arriving at the best fitting diagnostic category. Moreover, by appreciating the increasing
320 empirical basis for pleiotropy, clinicians can engage better in discussion with patients who
321 are often bewildered by the range of diagnoses they may receive across their lifespans.
322 Clinicians in other medical disciplines would not assign to a single clinical entity all the
323 physical ailments associated with a pleiotropic risk factor such as smoking, and there is no
324 reason why psychiatrists should either.

325
326 Children with congenital malformations, developmental delay, and ASD are already being
327 referred for molecular diagnostics, particularly for known pathogenic CNVs, but as the data
328 continue to accumulate, more types of genetic findings will be incorporated. It has been
329 argued that CNV testing should be offered to people with other forms of psychiatric
330 disorders; for now, the case is strongest for schizophrenia⁴⁰ but we predict ADHD is likely to
331 follow suit. The range of arguments for and against this are beyond the scope of this
332 article⁴¹; here, we note that identifying carriers of high penetrance mutations is currently of
333 limited value in psychiatry for precision medicine, but should testing be offered for
334 counseling or predictive purposes, it is important to consider the pleiotropic effects of
335 mutations. CNVs detected in children referred for testing may have important adult
336 psychiatric implications, and conversely if adults are tested, pleiotropy has implications for

337 their children and other relatives. The counseling challenges are substantial given the wide
338 range of possible outcomes, and much of the data that are required to do this with precision,
339 even for well-documented pathogenic CNVs, is lacking.

340
341 The extensive pleiotropy revealed by psychiatric genetics also has important implications for
342 interpreting mechanistic studies, whether in humans, using endophenotypes (Box 1), or in
343 animal and cellular models. Even for high penetrance alleles, the possibility of pleiotropy
344 implies the need for caution in ascribing a causal role in disease for particular brain imaging
345 correlates of that mutation, or in a rodent or stem cell model, neurobiological outcomes. This
346 issue has been discussed conceptually in the case of human endophenotypes and some of
347 the statistical approaches to identifying mediation outlined^{42,43}. The challenges to interpreting
348 results from model systems are more testing and will require researchers to cast the net
349 wider than is often the case in seeking the consequences of genetic risk factors and to relate
350 their findings to comparable findings from clinical neuroscience. This will require the use of
351 translatable measures and direct comparisons of the effects of genetic risk across levels of
352 complexity³.

353
354 Finally, on a positive note, pleiotropy may offer unsuspected therapeutic opportunities if it
355 turns out that this is reflected in shared pathophysiology. It is not uncommon for psychiatrists
356 to offer (off-label) treatments to patients with a particular diagnosis that are known to be
357 effective in a different psychiatric disorder. In a very general sense, pleiotropy can be seen
358 as offering some *post-hoc* justification for this, although we stress currently not at the level of
359 any specific treatment. As new treatments are developed to target one disorder, it is likely
360 that treatment will have a broader therapeutic role, and that wider patient populations may
361 benefit from advances in research into a particular disorder.

362

363

364

365

366 References

367

368 1 Whiteford, H. A. *et al.* Global burden of disease attributable to mental and substance
369 use disorders: findings from the Global Burden of Disease Study 2010. *Lancet* **382**, 1575 –
370 1586, doi: 10.1016/S0140-6736(13)61611-6 (2013)

371 2 Salomon, J. A. *et al.* Common values in assessing health outcomes from disease
372 and injury: disability weights measurement study for the Global Burden of Disease Study
373 2010. *Lancet* **380**, 2129-2143, doi: 10.1016/S0140-6736(12)61680-8 (2013)

374 3 Owen, M. J. New approaches to psychiatric diagnostic classification. *Neuron* **84**, 564-
375 571, doi:10.1016/j.neuron.2014.10.028 (2014).

376 4 Malaspina, D. *et al.* Schizoaffective Disorder in the DSM-5. *Schizophr Res* **150**, 21-
377 25, doi:10.1016/j.schres.2013.04.026 (2013).

378 5 van Os, J. & Kapur, S. Schizophrenia. *Lancet* **374**, 635-645, doi:10.1016/S0140-
379 6736(09)60995-8 (2009).

380 6 Lichtenstein, P. *et al.* Common genetic determinants of schizophrenia and bipolar
381 disorder in Swedish families: a population-based study. *Lancet* **373**, 234-239,
382 doi:10.1016/S0140-6736(09)60072-6 (2009).

383 7 Kapur, S., Phillips, A. G. & Insel, T. R. Why has it taken so long for biological
384 psychiatry to develop clinical tests and what to do about it? *Mol Psychiatr* **17**, 1174-1179,
385 doi:10.1038/mp.2012.105 (2012).

386 8 Linden, D. E. The challenges and promise of neuroimaging in psychiatry. *Neuron* **73**,
387 8-22, doi:10.1016/j.neuron.2011.12.014 (2012).

388 9 Sanders, S. J. *et al.* Insights into Autism Spectrum Disorder Genomic Architecture
389 and Biology from 71 Risk Loci. *Neuron* **87**, 1215-1233, doi:10.1016/j.neuron.2015.09.016
390 (2015).

391 10 Gaugler, T. *et al.* Most genetic risk for autism resides with common variation. *Nat*
392 *Genet* **46**, 881-885, doi:10.1038/ng.3039 (2014).

393 11 Schizophrenia Working Group of the Psychiatric Genomics, C. Biological insights
394 from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-427,
395 doi:10.1038/nature13595 (2014).

396 12 Rees, E. *et al.* Analysis of copy number variations at 15 schizophrenia-associated
397 loci. *Brit J Psychiat* **204**, 108-114, doi:10.1192/bjp.bp.113.131052 (2014).

398 13 Fromer, M. *et al.* De novo mutations in schizophrenia implicate synaptic networks.
399 *Nature* **506**, 179-184, doi:10.1038/nature12929 (2014).

400 14 Purcell, S. M. *et al.* A polygenic burden of rare disruptive mutations in schizophrenia.
401 *Nature* **506**, 185-190 (2014).

402 15 Singh, T. *et al.* Rare loss-of-function variants in SETD1A are associated with
403 schizophrenia and developmental disorders. *Nat Neurosci* **19**, 571-577, doi:10.1038/nn.4267
404 (2016).

405 16 Green, E. K. *et al.* Copy number variation in bipolar disorder. *Mol Psychiatr* **21**, 89-
406 93, doi:10.1038/mp.2014.174 (2016).

407 17 Hyde, C. L. *et al.* Identification of 15 genetic loci associated with risk of major
408 depression in individuals of European descent. *Nat Genet* **48**, 1031 – 1036,
409 doi: 10.1038/ng.3623

Formatted: German (Germany)

Formatted: Spanish (Spain)

Formatted: French (France)

410 18 Paaby, A. B. & Rockman, M. V. The many faces of pleiotropy. *Trends Genet* **29**, 66-
411 73, doi:10.1016/j.tig.2012.10.010 (2013).

Formatted: German (Germany)

412 19 Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M. & Smoller, J. W. Pleiotropy in
413 complex traits: challenges and strategies. *Nat Rev Genet* **14**, 483-495, doi:10.1038/nrg3461
414 (2013).

415 20 Kirov, G. *et al.* The penetrance of copy number variations for schizophrenia and
416 developmental delay. *Biol Psychiatry* **75**, 378-385, doi:10.1016/j.biopsych.2013.07.022
417 (2014).

418 21 Rees, E. Analysis of intellectual disability copy number variants implicates novel
419 schizophrenia risk loci. *JAMA Psychiatry* (2016; in press).

420 22 Rujescu, D. *et al.* Disruption of the neurexin 1 gene is associated with schizophrenia.
421 *Hum Mol Genet genetics* **18**, 988-996, doi:10.1093/hmg/ddn351 (2009).

422 23 Deciphering Developmental Disorders, S. Large-scale discovery of novel genetic
423 causes of developmental disorders. *Nature* **519**, 223-228, doi:10.1038/nature14135 (2015).

424 24 International Schizophrenia, C. *et al.* Common polygenic variation contributes to risk
425 of schizophrenia and bipolar disorder. *Nature* **460**, 748-752, doi:10.1038/nature08185
426 (2009).

427 25 Cross-Disorder Group of the Psychiatric Genomics, C. *et al.* Genetic relationship
428 between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* **45**, 984-
429 994, doi:10.1038/ng.2711 (2013).

430 26 Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and
431 traits. *Nat Genet* **47**, 1236-1241, doi:10.1038/ng.3406 (2015).

Formatted: French (France)

432 27 Hamshere, M. L. *et al.* Shared polygenic contribution between childhood attention-
433 deficit hyperactivity disorder and adult schizophrenia. *Brit J Psychiat* **203**, 107-111,
434 doi:10.1192/bjp.bp.112.117432 (2013).

Formatted: French (France)

435 28 Samocha, K. E. *et al.* A framework for the interpretation of de novo mutation in
436 human disease. *Nat Genet* **46**, 944-950, doi:10.1038/ng.3050 (2014).

437 29 Kirov, G. *et al.* De novo CNV analysis implicates specific abnormalities of
438 postsynaptic signalling complexes in the pathogenesis of schizophrenia. *Mol Psychiatry* **17**,
439 142-153, doi:10.1038/mp.2011.154 (2012).

440 30 Niarchou, M. *et al.* Psychopathology and cognition in children with 22q11.2 deletion
441 syndrome. *Brit J Psych* **204**, 46-54, doi:10.1192/bjp.bp.113.132324 (2014).

Formatted: French (France)

442 31 Stefansson, H. *et al.* CNVs conferring risk of autism or schizophrenia affect cognition
443 in controls. *Nature* **505**, 361-366, doi:10.1038/nature12818 (2014).

444 32 Tansey, K. E. *et al.* Common alleles contribute to schizophrenia in CNV carriers. *Mol*
445 *Psychiatr*, doi:10.1038/mp.2015.143 (2015).

446 33 Rees, E. *et al.* Evidence that duplications of 22q11.2 protect against schizophrenia.
447 *Mol Psychiatr* **19**, 37-40, doi:10.1038/mp.2013.156 (2014).

448 34 Lee, S. *et al.* Optimal unified approach for rare-variant association testing with
449 application to small-sample case-control whole-exome sequencing studies. *Am J Hum*
450 *Genet* **91**, 224-237, doi:10.1016/j.ajhg.2012.06.007 (2012).

451 35 Ruderfer, D. M. *et al.* A family-based study of common polygenic variation and risk of
452 schizophrenia. *Mol Psychiatr* **16**, 887-888, doi:10.1038/mp.2011.34 (2011).

453 36 Hamshere, M. L. *et al.* Genome-wide significant associations in schizophrenia to
454 ITIH3/4, CACNA1C and SDCCAG8, and extensive replication of associations reported by
455 the Schizophrenia PGC. *Mol Psychiatr* **18**, 708-712, doi:10.1038/mp.2012.67 (2013).

456 37 Flint, J., Timpson, N. & Munafò, M. Assessing the utility of intermediate phenotypes
457 for genetic mapping of psychiatric disease. *Trends Neurosci* **37**, 733-741,
458 doi:10.1016/j.tins.2014.08.007 (2014).

459 38 Franke, B. *et al.* Genetic influences on schizophrenia and subcortical brain volumes:
460 large-scale proof of concept. *Nat Neurosci*, doi:10.1038/nn.4228 (2016).

461 39 Cuthbert, B. N. The RDoC framework: facilitating transition from ICD/DSM to
462 dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry*
463 **13**, 28-35, doi:10.1002/wps.20087 (2014).

464 40 Costain, G. *et al.* Pathogenic rare copy number variants in community-based
465 schizophrenia suggest a potential role for clinical microarrays. *Hum Mol Genet* **22**, 4485-
466 4501 doi: 10.1093/hmg/ddt297 (2013).

467 41 Gershon, E. S. & Alliey-Rodriguez, N. New ethical issues for genetic counseling in
468 common mental disorders. *Am J Psychiat* **170**, 968-976,
469 doi:10.1176/appi.ajp.2013.12121558 (2013).

470 42 Walters, J. T. R. & Owen, M. J. Endophenotypes in psychiatric genetics.
471 *MolPsychiatr* **12**, 886-890, doi:10.1038/sj.mp.4002068 (2007).

472 43 Kendler K. S. & Neale, M. C. Endophenotype: a conceptual analysis. *MolPsychiatr*
473 **15**, 789-97, doi: 10.1038/mp.2010.8 (2010).

474 44. Antilla, V. *et al.* Analysis of shared heritability in common disorders of the brain.
475 Preprint at *bioRxiv* <http://dx.doi.org/10.1101/048991> (2016).

476

Formatted: German (Germany)

Formatted: German (Germany)

477 Figure Legends

478 Figure 1. Types of pleiotropy.

479

480 Adjacent genes containing functional variants (FV; yellow circles) that together directly
481 influence four distinct phenotypes (blue shapes). The three phenotypes directly influenced
482 by FV₁ and FV₂ are examples of genic pleiotropy. The circle and pentagon phenotypes
483 influenced by FV₁ are examples of allelic pleiotropy in which the same variant rather than
484 just the same gene influences multiple phenotypes. FV₁ influences the triangle phenotype
485 but this is indirect and only occurs through the direct effects of FV₁ on the circle phenotype.
486 This is mediated pleiotropy. Alleles at FV₂ and FV₃ are correlated (through linkage
487 disequilibrium; LD) with the same single nucleotide polymorphism (SNP) Accordingly, that
488 SNP will be associated with both phenotypes that are caused by those functional variants.
489 The SNP is depicted midpoint between the genes but be positioned anywhere within the
490 region of LD, including within one of the genes. This is pseudo pleiotropy due to co-
491 localization. This region is also prone to a deletion CNV which results in complete loss of
492 function of both genes by virtue of which it is associated with all five phenotypes. In a literal
493 sense, all of the blue phenotypes in this instance are now examples of allelic pleiotropy
494 (being directly caused by the same CNV allele at a single locus).

495

496

497 Figure 2. Relative CNV Frequencies.

498

499 Relative frequencies for schizophrenia associated CNVs. Frequency is expressed as fold
500 increase in each disorder relative to the estimated population frequency. Data are taken
501 from²¹ based on loci reported as schizophrenia associated^{12,21}. CNVs are described by
502 cytogenetic position or the named syndrome most strongly affiliated with the locus. The
503 approximate lifetime population risk for SZ is approximately 1% and for ID/ASD combined
504 4%²⁰. Abbreviations: SZ schizophrenia; ASD/ID autism spectrum disorder and ID intellectual
505 disability combined. WBS Williams-Beuren Syndrome. PWS/AS. Prader-Willi
506 Syndrome/Angelman Syndrome; VCFS Velo-cardio-facial syndrome; CNV copy number
507 variant; del deletion; dup duplication.

508

509 Figure 3. Genetic correlation between schizophrenia and selected psychiatric disorders.

510

511 Psychiatric disorders showing significant evidence ($P \leq 0.001$) for overlaps between common
512 variant contributions to schizophrenia and other psychiatric disorders. Overlaps are

513 expressed as correlation in heritability (r_g) captured by SNPs. Data are from⁴⁴. Abbreviations:
514 ADHD Attention Deficit Hyperactivity Disorder; OCD Obsessive Compulsive Disorder.

515

516 **Box 1: Patient Stratification**

517 There is general agreement that we need new approaches to patient stratification in
518 research if we are to better understand gene-phenotype relationships, accelerate
519 understanding of aetiology and pathogenesis, and inform mechanistic studies and treatment
520 trials. Generally speaking three ways in which we can move beyond the constraints of
521 current diagnostic approaches have been proposed. Rather than being mutually exclusive,
522 these can be thought of as targeting psychiatric disorders at different levels of conceptual,
523 and aetiological complexity, from the molecular at one end to the function and behaviour of
524 the whole human at the other. Models that attempt to capture this hierarchical complexity of
525 have been proposed and discussed in detail elsewhere³.

526

527 First, we can use clinical classifications that cut across or divide current diagnostic groups.
528 These might be based upon the presence of absence of particular symptoms (e.g.
529 hallucinations), syndromes (e.g. psychosis, depression) or other features such as course or
530 outcome. This may aid the identification of risk factors and pathogenic mechanisms
531 providing the strata map more closely onto these than do current diagnostic groupings. This
532 approach also has the potential to help our understanding of the basis of heterogeneity.
533 There is some evidence to support this type of approach, for example stratifying people with
534 BD for the presence of psychotic symptoms predicts a higher burden of schizophrenia risk
535 alleles, and, conversely, stratifying people with schizophrenia for presence of manic type
536 symptoms predicts a higher burden of bipolar risk alleles^{35,36}. These preliminary findings
537 suggest that, across disorders, sets of syndromes have some shared biological basis, and
538 support a model where disorders, as manifest in individuals, may be viewed as the
539 confluence of partly orthogonal symptom dimensions.

540

541 Second, stratification can be based on the presence of a particular aetiological factor (e.g. a
542 rare high penetrance mutation, a particular environmental exposure) rather than clinical
543 features. The assumption is that constraining the risk architecture will increase biological
544 homogeneity, and allow researchers to focus on specific risk mechanisms and understand
545 what factors lead to different outcomes, including resilience as well as risk. This type of
546 approach also lends itself to complementary studies in cells and animals as well as humans.
547 In psychiatry, this has yet to yield unqualified success, and, given evolutionary multi
548 purposing of proteins, which may have different functions in different cells or cell
549 compartments, even a single genetic variant might map onto different pathogenic

550 mechanisms in carriers. While this is a theoretical concern, the fact that regardless of the
551 specific psychiatric diagnoses (ID, ASD, schizophrenia), rare *de novo* and LoF mutations
552 tend to impact upon similar broadly similar processes (e.g. glutamatergic pathways
553 regulating synaptic plasticity, chromatin modifiers, and targets of fragile X mental retardation
554 protein) suggests that individual mutations are likely to influence the same pathogenic
555 mechanisms across disorders¹³ .

556
557 Third, in attempting to relate risk factors and clinical phenotypes to underlying
558 pathophysiology and mechanisms, stratification can be performed at the level of
559 endophenotypes (intermediate phenotypes). One problem with this approach is the large
560 number of potential endophenotypes including measures of cognition, brain structure,
561 electrophysiology, and biochemistry. Moreover, initial claims that endophenotypes are likely
562 to be less complex genetically than clinical disorders have not in general been supported³⁷
563 and perhaps this explains why failures to link endophenotypes to genetic risk³⁸ are for now
564 more notable than any reproducible successes. Nevertheless, this approach offers a means
565 by which genetic risk can be linked to disturbances of brain function, and a framework for
566 doing so has been implemented in the Research Domain Criteria (RDoC) project of the
567 National Institutes of Mental Health³⁹. The pleiotropic effects of many risk alleles are clear
568 reminders that there are pitfalls associated with using this approach to chart the pathways
569 mediating the effects of genetic risk on clinical phenotypes (see main text).

570
571
572
573
574