Department for Education

Genomics: Implications for education

A futures report

IPSOS UK
Strategy and Advisory

Authors: Joel Hooper, Marzieh Azarbadegan, Evie Cogley, Michelle Mackie, Nathan Bransden

September 2024

Department for Education working with Government Office for Science

This research was commissioned by the previous government and is not necessarily a reflection of the current government's policies and priorities. DfE is publishing this report in the interests of transparency.

Contents

Introduction

In January 2024, building on the Genomics Beyond Health report, the Department for Education (DfE), with co-funding from the Government Office for Science (GO-Science), commissioned Ipsos UK through the Futures Procurement Framework to understand the potential future risks and opportunities of the use of genomics in education. The Government Office for Science's Beyond Health report (Government Office for Science, 2022) highlighted two potential issues to be explored. Firstly, the potential expansion of an unregulated commercial market in genomic testing in education-related fields. Secondly, the potential use of genomic screening at birth to identify additional educational needs before other data is available; for example, to identify children at higher risk of developing conditions associated with Special Educational Needs and Disability (SEND).

To note, the two scenarios detailed above and presented within this report are intended to explore two potential future outcomes, and do not constitute government policy or government recommendations. They serve as projections based on specific assumptions and should not be interpreted as definitive forecasts or government recommendations.

As part of this project, Ipsos reviewed the most relevant literature, held a workshop with key relevant stakeholders in the field, and discussed findings with relevant policy teams in DfE to discuss emerging themes and concerns. **This summary paper explores the policy implications of the two potential future scenarios and makes recommendations to the DfE for next steps.** These recommendations are not government recommendations, but an output of this research project. They draw from the detailed project outputs, which include: a rapid evidence review (Annex A), future pathways (Annex B), and a short summary of the commercial market (Annex C).

What is genomics?

Genomics is a field of science that studies the structure, function, and evolution of genomes, which are the complete set of genes (segments of DNA) present in a cell or organism. This field includes a variety of scientific disciplines, such as genome sequencing (identifying different genetic variations and mutations) and bioinformatics (the interpretation of large biological datasets). In relation to education, genomic technologies can be used to analyse how genes influence traits and conditions that impact on educational outcomes, such as intelligence, educational attainment, and SEND. This is a highly complex area; in this summary paper Ipsos repeat the following key terms throughout defined as:

- **Genome:** The complete set of genes or genetic material present in a cell or organism that contains all of the information needed to build and maintain that organism.
- **Heritability**: the extent to which genetic differences among individuals contribute to observable characteristics or phenotypes.
- **Trait**: an observable characteristic or behaviour.
- **Conditions**: physical and developmental conditions associated with SEND.
- **Educational outcomes**: the full range of potential measures of academic achievement as well as duration of time in education (which is used in studies to describe educational attainment).
- **Phenotype:** observable physical traits (from visible traits such as hair colour, to more complex traits such as behavioural patterns). These are the result of the interaction between our genes (a genotype refers to the specific set of genes related to a particular phenotype) and the environment.
- **Genomic screening:** refers to examining someone's genome to identify variations, mutations or anomalies.
- **Monogenic:** Traits that are the result of a single gene.
- **Polygenic:** Traits that are the result of multiple genes.

How Genomic technologies relate to education

Genomic usage currently: The NHS has committed to screening 500,000 whole genomes by 2023/2024 to help speed up diagnosis and improve outcomes (NHS England). There is also a global Direct-to-Consumer Genetic Testing Market (DCGT) for products that provide information on a range of topics, from ancestry to personalised healthcare and wellness advice based on a customer's genetic information. Prominent companies include (but are not limited to) 23andMe and AncestryDNA. The market is growing but is concentrated in the US.

Polygenic Scoring (PGS): Polygenic Scoring (PGS) is a statistical measure of the extent to which genomic variations contribute to particular traits, conditions and outcomes. PGS is calculated by researchers first identifying genomic variants associated with complex diseases, through comparing the genomes of individuals with and without those diseases. The enormous amount of genomic data now available enables researchers to calculate which variants tend to be found more frequently in groups of people with a given disease. This information is then processed using a statistical model on a computer to estimate how the collection of an individual's variants affects their risk for a certain disease. This yields a polygenic (risk) score or PGS. These scores are based on large population studies; the most developed PGS is called Eduyears/EA3. From analysis of more than a thousand genetic variants from a sample of 1.1 million participants, Eduyears explains 11–13% of differences between people in how long they stay in education (Lee, et al., 2018), and 7–10% of differences in cognitive ability (Allegrini, et al., 2019). PGS applies to measures of educational attainment; this builds statistical associations between measures like academic performance, intelligence quotient (IQ), duration in school and specific genetic structures.

The heritability of a condition is indicative of the potential for genomic technologies to identify the genetic structure of the condition within the genome. Autism Spectrum Conditions (ASC) is estimated to be up to 91% heritable (Zhou, et al., 2023), and screening for variations identified in population studies as 'strongly associated with ASC' have been used to detect these variations in individuals. Ipsos have highlighted the evidence on ASC, ADHD and dyslexia (see Table 1 in Annex A). These conditions overlap and research indicates that understanding the genetic composition of ASC, also contributes to better understanding of Attention Deficit Hyperactivity Disorder (ADHD), dyslexia and other conditions.

Nature vs Nurture: For many traits, conditions and outcomes, it is difficult to unpick the role of genes versus the environment children are brought up in. Children can inherit traits from their parents through biological and non-biological means. For instance, if a child is brought up in a house full of books, and they later go on to be literate at a younger age than most, it is difficult to unpick how influential the environment was over their genetic propensity for literacy. Additionally, PGS scores are based on large population studies and so are a less accurate predictor of individual outcomes.

The Potential Utility of Genomics in Education

Prior educational attainment and teacher observation will remain far more accurate than PGS scores for predicting an individual child's educational attainment, cognitive ability, and IQ. Currently, research indicates that 7-10% of differences in cognitive ability can be predicted using PGS scores. The literature and experts both anticipate that at most 10-20% of differences in educational attainment will be predicted using genomic screening in the near future. In addition, this predictive power does not operate at an individual level and false positives will occur. Therefore, information on an individual's prior educational attainment and teachers' understanding of an individual child is more accurate, reliable and cost-efficient (Martschenko, et al., 2019) in the current state of knowledge – and likely to be so in the near future.

However, there is some potential to use Polygenic Scores to identify relevant genetic variations related to specific conditions. For example, a large population-based study of 20,000 people with autism calculated the PGS for autism and found that people with the highest scores are nearly three times as likely to have autism as those with the lowest scores (Grove, et al., 2019). Although, PGS has more efficacy and predictive power within sub-groups of people who already have an identified increased baseline risk, in contrast to the general population where PGS cannot be used to predict risk for autism accurately. For example, PGS would have stronger predictive power when applied to a sub-group of people who have an autism-related mutation, which can be identified through other genetic screening mechanisms. This is because a high PGS within this sub-group would provide a stronger indication of risk for developing autism (beyond the baseline risk already highlighted). This is an important aspect of why the literature indicates more promising capabilities for this use of genomics, related to screening for conditions, over the use of PGS to predict educational outcomes.

Screening for specific conditions: There is also a large body of literature on the application of genomics to understand the genetic components of developmental conditions. This research includes PGS amongst other mechanisms for assessing genetic associations and risks. The literature covers a diverse range of developmental conditions from Autistic Spectrum Conditions[1] (ASC), to Attention Deficit Hyperactivity Disorder (ADHD), and developmental dyslexia. Conditions are caused by variations in the genome that can be polygenetic and monogenic in nature, meaning that there is not one single approach to understanding the genetic causes of conditions. For instance, certain combinations of common variants increase the likelihood of ASC in people with rare, inherited mutations linked to the condition (Weiner, et al., 2017).

Therefore, genomics could provide useful information on specific conditions before the availability of other credible data, especially in the early years of a child's life, where other forms of assessments for such conditions cannot be done. Screening could provide more information to parents, most likely in an early years context, which

[1] Healthcare literature sometimes uses 'disorder' rather than 'condition'

could be used to adapt the home learning environment and childcare provision, begin educational planning for starting school (if/when appropriate), and allow for streamlining of children at elevated risk for the associated assessment/diagnosis (Martschenko, et al., 2019).

Ipsos' assessment is that within the relatively near future, there will be sufficient scientific and technological developments to perform screenings like this en masse. Currently, for example, Chromosomal Microarray Analysis is a commonly ordered screening tool for ASC and looks at chromosomes to see if there are extra or missing parts that could cause ASC. This screening mechanism, alongside others which analyse different aspects and variations in the genome associated with ASC, provide an example of what future genomic screening technologies could look like. Most likely there will not be a singular screening mechanism but multiple that cover different relevant features of the genome associated with a condition. Moreover, the utility of genomics depends on the condition. For example, it could be used to identify a heightened propensity to develop ASC from an early age. However, for ADHD, which research suggests may be more subject to environmental factors over a person's youth to adulthood, this application of genomics is likely to be less useful. In such a case, some literature indicates the utility of sub-field of genomics (pharmacogenomics) to assess genetic propensities for certain medications (Balogh, et al., 2022).

Larger and more diverse sample sizes are required. For genomic technologies to have a significant degree of accuracy and applicability, larger sample sizes for whole genome sequencing and genetic studies would be required so that findings made thus far can be replicated amongst the larger population. Much more representative and diverse samples for Whole Genome Sequencing (WGS) and genetic studies are required so that findings can be replicated outside of people of European ancestry (Rabinowitz, et al., 2019).

What the future might look like (see Annex B for full pathways)

As part of this work, Ipsos developed and tested scenarios with a group of experts. These were then refined into two pathways through which improvements in genomics technologies and more screening could develop and have an impact on education. The first, called the 'Professional' pathway, relates to genomic screening and testing in the healthcare system for specific conditions. The second, called the 'Consumer' pathway, refers to the development of products specifically marketed for education by Direct-to-Consumer Testing Companies (DTCT) such as 23andMe. It is important to note that these are potential pathways based on current evidence and Ipsos' engagement with experts, rather than predictions.

'Professional'

The NHS is screening an increasing number of babies and children. The UK is a leader in Genomics; the NHS already screens ill babies (neonates) and is aiming to be the first national healthcare system to offer whole genome sequencing as part of routine care. As more babies and young children are screened and more scientific research improves our understanding of certain conditions (e.g. ASC or ADHD), a greater number of healthcare practitioners and parents or care-providers could have access to genetic data. This is already happening and is likely to continue.

As more parents and caregivers have access to genomic information, they may start using this information as evidence for the need for additional provision or SEND. This could lead to an increased demand for private genomic screening relating to the risk of certain SEND conditions. In this case, more affluent parents or care-providers could afford to benefit from this whilst less affluent parents would not. Additionally, it is unclear how schools would utilise genomic information appropriately if provided through private means, outside of the legitimacy of the NHS or government bodies. There is little evidence of this happening yet in the UK, but this appears to occur in the USA currently to some extent. Ipsos' assessment is that it's likely to happen in the UK over the next 5-10 years to assess risk of certain conditions like ASC from an earlier age.

'Consumer'

Direct-to-Consumer Genetic Testing (DTGT) Companies could start marketing genomic tests which promise to identify someone's genetic pre-disposition to higher educational attainment, this could range from general cognitive ability to mathematical ability, and even IQ. This could lead to parents and children using tests to inform their choices and to expect schools to act on this information. DTGT companies may be more likely to focus more heavily on 'Consumer' genomics because they will not have to offer legal disclaimers relating to clinical issues. Based on current regulation, these could be marketed in the same way as tests for ancestry (for example, 23&Me). Stakeholders

consulted had varied views on this, with some advising that a market similar to the one in the USA may develop in the UK, and others flagging that the predictive power of the tests are too limited for companies to market and risk losing their reputation. The market analysis conducted by Ipsos also signifies that the market for genomics tests in the UK is not very big. It is, therefore, unlikely that commercial companies will start advertising 'Consumer' tests until they get to a certain level of accuracy, as maintaining a credible reputation is important. However, this could change depending on technological development and public demand. In this case, more affluent parents or care-providers could afford to benefit from this whilst less affluent parents would not.

Core issues

This project did not explore the regulation, standards and guidance which applies to genomics in detail. There is no overall regulatory framework of the private market at present, and the project did not explore the detailed regulation that may be relevant from different regulators; the NHS sets the policies and procedures for the applications of genomics within its own genomics services and provides guidance for those seeking genomic services through DTGT companies. However, Ipsos identified the following as key issues the government may need to address, including, but not limited to, through regulation.

Public perceptions. Research shows that the public is likely to perceive the use of genetics as 'deterministic'. Any discussion of genomics could be perceived as an attempt to pre-determine the outcomes for children. The use of genetic data in education could inadvertently limit pupils' opportunities and growth because they feel they are unlikely to succeed in a particular subject or area, and are therefore less likely to try to develop. This could potentially lead to discrimination and stigma, including self-stigma, particularly for young people who are still developing.

Privacy. There are three broad concerns regarding privacy. First, anonymisation of personal genomics data is particularly difficult compared to other forms of data, since genetic information is shared between relatives and ethnic groups, making it possible to narrow down the identity of a person even in an anonymised database (Lewis, et al., 2020). Second, there are concerns about sharing and managing large-scale datasets for research projects that are run internationally that respect national boundaries (Birney, et al., 2017). Third, in the education setting (vs. clinical setting), it is much less clear who has the right to access and use students' data.

Bias in the data. Those who are not of European ancestry are under-represented in current genomic studies. This means PGS scores for these groups are likely to be less accurate and cannot be utilised without accounting for their existing underrepresentation in the datasets (Rabinowitz, et al., 2019). Furthermore, as genomics relies on large data processing capabilities, the development of artificial intelligence (AI) could help in speeding up data processing, which in turn, could quicken the capability of genomics in

presenting more personalised information. However, the use of AI could exacerbate existing issues with the data and technology, such as algorithmic bias where a computer system systematically favours certain outcomes or groups due to errors in its data, design or usage.

Inequality. If DTGT companies commercialise genetic tests, there is a risk that only people from higher socioeconomic groups would access the tests, widening the inequality gap for those who are unable to access these tests and the relevant support by having knowledge of specific risk factors (Martschenko, et al., 2019).

Interactions with the Department for Education's current policy thinking

As part of this project, Ipsos and the Department for Education's Science Team engaged small groups of policy officials to think through some of the potential implications of this evolving field. In line with our findings on the potential utility of genomics, the key interaction is between the potential for genomic screening and SEND policy.

Genomic screening could function within current systems for SEND and alternative provision (AP) in the education system. For instance, genomic screening aligns with the current social model of disability, not being used to diagnose a child or offer a direct treatment, but rather to provide more information on a child with likely special needs to their family, early years setting, GP and local authority. It is possible that genomics could be used as a form of pre-diagnosis or assessment of risk for developing SEND, which could potentially improve systems of diagnosis, support, and early intervention. Understanding the potential risk of developing certain SEND conditions earlier could also mean developing an Education, Health, and Care plan (EHC) sooner for young people. This aligns with the social model of disability by understanding the potential needs a young person may have from an early stage and setting out the additional support needed to meet these needs. However, it is likely that this would be perceived as a potential medicalisation of the model which would need to be explained.

For example, most children with autism are not diagnosed until after the age of 3, partly due to the limitations of assessment systems and the ways autism develops in children. It is possible that future genomic screening could develop to the point where a child could be assessed for increased risk of developing autism from birth. Autism has a high level of heritability (Zhou, et al., 2023), more than many other conditions, and there are some indications that such screening could be achieved with sufficient accuracy in the next decade. With such information, early intervention and additional support could be put in place to both improve outcomes for the child entering the education system and the support provided to their family. It is also the case that early intervention could lessen the severity of autism in a child and improve long term life outcomes in-and-outside education (Johnson, et al., 2013). A 2021 study showed that children who are diagnosed with autism before age 2.5 show greater gains in their social skills, on average, than children who are diagnosed later. This improvement was associated with the ability of parents/relevant parties to increase the frequency of interventions during key early developmental stages (Gabbay-Dizdar, et al., 2022). Although using genomics to screen for such conditions will not provide a formal diagnosis, genomic information would still provide actionable data for parents, early years settings, GPs and Local Authorities that could significantly improve life outcomes.

The recently published Department for Education SEND and alternate provision green paper (HM Government, 2022) highlighted key issues for parents: inconsistent access to SEND support, overreliance on Education Health Care Plans (EHCP) for support when

they are hard to obtain, prolonged waiting times for diagnosis, and limited capacity for staff to deliver early intervention/system work to improve outcomes. Genomic testing could help address some of these perceived issues: by identifying a child's special needs earlier, improving parents' understanding of their child, streamlining support services for children at increased risk, and potentially allowing for staff to spend less time on diagnoses and assessments over systemic work. Systemic work relates to improving the learning environment and the effectiveness of a school's Special Educational Needs (SEN) and behaviour support systems. A recent DfE paper found that Educational Psychologists have significantly less time to focus on systemic work in schools due to individual assessments and diagnostic work taking up most of their time (Department for Education, 2023).

Conclusions and Recommendations

To note, the scenarios detailed above and presented within this report are intended to explore two potential future outcomes, and do not constitute government policy or government recommendations. They serve as projections based on specific assumptions and should not be interpreted as definitive forecasts or government recommendations.

In summary, and building on the future pathways (see p9-12 and Annex B):

1. If current trends continue, a substantial number of babies will receive a genetic screening over the next 5 years. Therefore, more and more parents or care-providers will have access to genetic information that could be associated with education, but does not directly provide information on cognitive, emotional, or social ability or risk markers for SEND conditions.
2. 'Consumer' genomics currently has little utility within education settings, but increased screening and marketing of education-related genomic products is anticipated in the relatively near future.
3. It seems likely that there will be growing public awareness of genomics in healthcare, and this is likely to start to be linked to educational outcomes by the public.

Based on this report, Ipsos recommends the Department for Education:

- **Monitor developments, in:**
    - Scientific research. A number of significant UK studies are underway that could impact genomics capabilities in the near future. Most of this work is being done in association with the NHS and government funding. Such as:
        - Generation Study / Newborn Genomes Programme. This will sequence the genomes of 100,000 newborn babies by 2025 (Genomics England).
        - UK Biobank have sequenced whole genomes of 500,000 volunteers. Combining it with their existing de-identified data from health records, lifestyle information, amongst other data points (UK BioBank).
        - NHS Long term Plan on Genomics: To sequence 500,000 whole genomes by 2024 and help transform healthcare for maximum patient benefit, including for all children with cancer or children who are seriously ill with a likely genetic disorder.
        - There are also rapid developments in the availability of genomic data for current UK national cohorts (e.g., the Millennium Cohort Study) - with linked genomic data planned to be available from late 2024. (NHS England).
    - The cost and forms of genomic screening.
        - Product launches and advertising from commercial companies.
        - The cost of genomic screening in the NHS.

- Any indication of genomic information being used in the educational system.

- **Could support the development of research related to genomics and education**
  - The NHS Long-Term Plan for Genomics aims to be the first health care system to utilise genomics, and it is already developing capacity in this space for clinical diagnostic purposes. Further research is required to explore many aspects of genomics within the SEND and AP space. As the standard for which diseases are screened is incredibly high in the NHS, with requirements for highly reliable tests and availability of follow-up treatment amongst other things, the NHS is unlikely to start screening for these conditions soon.
  - If the Department for Education wants to explore the potential use of genomic data in education, or feels its application is likely and wants to steward the system, the key priority should be to support scientific research and development, with an immediate focus on encouraging further studies on the link between genomics and SEND conditions. This support could include being involved in scientific research in this area, funding it, or signalling interest to other funding bodies, such as UKRI.
  - If there is a significant development in this space, the government will have to consider how this could impact early assessment and intervention. Schools, educators, parents or care-providers and young people will all need to have increased understanding of the potential utility and limitation of genomics data, otherwise, there is a risk of viewing the data deterministically and in turn, stigmatisation and labelling for young people.

If genomics technologies were utilised in education, Ipsos have identified the following as key policy considerations and questions in the shorter term:

- Would genomic screening for some subgroups or all of the population in early years education meet National Screening Committee criteria for effective and ethical screening? (UK National Screening Committee, 2024)
- Would screening at a younger age cause over / underdiagnosis in practice? For example, screening could be trialled and compared with the current diagnostic tests. Given some tests for SEND already exist, this could be done in the short term.
- If young people are being screened for SEND conditions, are there systems in place to support them with the appropriate interventions? Would screening lead to an increased number of children requiring assessments and provision which the system cannot provide?
- A public dialogue on Genomics could help the government to understand the public perception and manage the narrative (especially thinking about the worries of viewing genomics data as deterministic).

Longer term considerations and questions include:

- Do earlier interventions make a significant difference in the life outcomes of people (vs. if needs are identified as usual within the educational system)? This would also relate to the system's ability to respond with additional support.
- What would the uptake be? How many parents or care-providers would use this screening? And what would determine whether they use it or not? For example, the price of tests, how the accuracy is marketed, whether they are already worried about the heritability of specific conditions existing in their families. And if they use it, what kind of information would they need other than basic risk scoring?
- What kinds of support would educators/schools/teachers need if genomic data was used to identify SEND pupils? How would it differ for mainstream vs. SEND schools?

Annex A: Summary of evidence review

Introduction

As part of this project, Ipsos reviewed the most relevant literature, held a workshop with key stakeholders in the field, and discussed findings with relevant policy teams in DfE to highlight emerging themes and concerns. This document outlines the summary of a rapid evidence review (RER), covering 33 pieces of literature. For the RER, Ipsos used search terms agreed with the DfE (such as 'Genomics in education', 'Behavioural genetics', and 'DNA-based education') to find 46 pieces of relevant literature from 2018 onwards. Following this, Ipsos selected 25 pieces of literature that were most relevant to the themes and summarised the main findings in this document. A further eight pieces of literature were also explored to gather a more in-depth understanding of some of the topics.

The scientific possibilities and limitations of genomics/ genetic testing in education

What is genomics, and how has it been applied in education-related research

Genomics is a field of science that focuses on studying the structure, function, evolution, and mapping of genomes, which are the complete set of genes or genetic material present in a cell or organism. This field includes areas like gene sequencing and bioinformatics to analyse the function and structure of genomes.

Statistical analysis of large genetic data sets from many individuals assesses which combinations of genetic variations contribute to a 'trait' or condition, represented by a 'Polygenic Score' (PGS). PGS scores can be applied to educational outcomes and more accurately indicate the risk of certain neurodevelopmental conditions, such as autism spectrum condition (ASC), ADHD and dyslexia**,** and mental/psychological issues.

The capabilities of polygenic scores (PGS) to explain differences in educational attainment are improving quickly, even when controlling for factors such as socioeconomic status. To briefly summarise the scientific developments in this field, the most powerful polygenic score to date (as of early 2024) is known as EduYears or EA3. This polygenic score is made up of more than 1,000 genetic variants and was derived from a sample of 1.1 million participants (Lee, et al., 2018). It can explain 11–13% of differences between people in how long they stay in education and 7–10% of differences in cognitive ability / cognitive traits (Allegrini, et al., 2019). One high-level UK-based study shows that at age 16, polygenic scores (PGS) account for 14% of the variance in educational attainment, compared to 23% for their parent's socioeconomic status

(Selzam, et al., 2016). This indicates that further advancements in this field might allow for accurate prediction of educational attainment within the variance related to PGS.

Genomics can also be used to assess the risk of certain neurodevelopmental conditions such as autism spectrum disorder (ASC), attention deficit hyperactivity disorder (ADHD) and dyslexia (Martschenko, et al., 2019). Importantly, the assessment of elevated risk for such conditions is highly complex, and it is agreed that there is a complex interplay between genetic risk for such conditions and environmental factors. This means that genomic screening could never provide a diagnosis per se, but rather highlight an elevated risk for certain conditions based on genetic data. Outside of these conditions, the literature around the use of genomics to assess the risk of mental and behavioural disorders is more nuanced. Although, there does appear to be some promising research using PGS to predict mental health outcomes (Anderson, et al., 2019) as well as assess risk for neuropsychiatric conditions (Jansen, et al., 2018). Generally, this research is less developed than genomics research on educational attainment and neurodevelopmental conditions. Therefore, further research on the application of genomics to mental/behavioural conditions would be required to understand this properly. The table below summarises key findings related to specific conditions (Table 1).

Table 1: Summary of current research on specific conditions

SEND Condition	Genomics Research	Potential Impact
Autism Spectrum Disorder	ASC is a polygenetic condition; it is caused by a complex interaction of gene variations with environmental factors. The estimated heritability of ASD is up to 91%. There are a variety of screening procedures for ASC from Karyotype analysis, fluorescence in-situ genomic hybridization (FISH), fragile X testing, chromosome microarray analysis (CMA), and whole genome/exome sequencing (WGS/WES). One genetic test called chromosome microarray analysis identifies certain genetic variations which are associated with autism has a detection rate of up to 24% of those tested (otherwise called the diagnostic yield). And another Whole Exome Sequencing has a diagnostic yield of up to 15% (Zhou, et al., 2023) Genomic testing shows significant promise for improving earlier diagnosis of different genetic components causing ASD to arise. This is likely to be replicated for other highly heritable conditions. The conditions overlap in the science – studies on ASD have led to findings about ADHD for example.	Most children with ASC are diagnosed after 36 months. This could provide a means to earlier identification and support for a child at higher risk for developing ASC. Not a diagnosis but is more information for parents/carers on their child that can inform early years support. Could cause stigmatisation of children without propre communication/ genetic counselling.
ADHD	The heritability of ADHD is estimated around 80%, relatively high. However, the potential genetic effects on risk, symptom severity, and persistence are less clear. ADHD seems more related to epigenetic influences on brain development. Genomics has highlighted increased risk of developing ADHD alongside related conditions. More research on the genetic architecture of ADHD is required (Balogh, et al., 2022)	It is possible to manage ADHD in a number of ways though genetic testing beyond assessing risk. Specifically, through pharmacogenomic testing clinicians can use this information to more accurately determine whether certain medications are likely to work[38].

SEND Condition	Genomics Research	Potential Impact
Dyslexia	Developmental dyslexia (DD) is a learning disorder affecting the ability to read, with a heritability of 40–60%. Whole genome sequencing has revealed more about the genetic architecture of DD, relating it to the literature on PGS for educational attainment (Gialluisi, et al., 2020)	Through genetic testing, genetic predisposition for developing Dyslexia could be better understood. Understanding genetic causes of DD also sheds lights on other conditions

Limitations

There are four significant limitations of PGS currently (Asbury, et al., 2021):

- PGS are reliably predictive at the group level but not at an individual level;
- the heritability of a trait varies in different environmental contexts;
- genes can influence educational attainment even when they are not passed on biologically through the rearing environment (akin to how parents' genes affect their children's behaviour); and,
- the sample for PGS is still very limited to genetically European populations.

The first three limitations highlight the (current) fundamental restrictions of genomics to accurately predict an individual's educational performance, making genetic testing for measures such as IQ highly unreliable and scientifically unsubstantiated. However, the literature suggests that in the next two decades, an individual could be assessed for their genetic propensity for some educational outcomes, neurodevelopmental conditions, other relevant SEND-related issues relative to their home life, socio-economic status and other factors. There will most likely never be a high level of certainty (Martschenko, et al., 2019), and the predictive power of such tests on educational attainment will produce false positives.

The last major limitation relates to the under-representation of people of non-European ancestry, especially African ancestry. One paper specifically focuses on understanding how PGS for educational attainment maps onto a vastly underrepresented group in the data: African Americans (Rabinowitz, et al., 2019). It finds that there is still explanatory power of PGS in this demographic, as high PGS still equates to high educational attainment, although much further research is required. It is also true that different ethnicities can carry different neurodevelopmental traits, which could lead to an underdiagnosis in underrepresented groups (Rabinowitz, et al., 2019). Without much consideration of addressing the issue of limited variance in ethnicities covered in genomics research, the application of genomics to educational settings would be practically useless.

Some argue that if polygenic scores are likely to have value, it will be in the context of very early intervention during the preschool years before other types of data (such as prior attainment) become available. This is due to the existence of other measures that are often more predictive and don't share the same ethical concerns as PGS. For instance, prior educational attainment captures genetic influences and tells us much more about educational performance in children and young people than PGS/EduYears scores.

Current applications of genomic screening in relation to education

The potential market for genomics in education is complex and highly dependent on government policy/regulation in the next decade. The literature suggests that there are potentially two routes through which genomics could be of potential value in educational contexts: Clinical / 'Health' applications and 'Consumer' services.

Clinical / 'Health' / 'Professional' genomics

Clinical genomics relates to the use of Polygenetic Scores (PGS) to assess well-evidenced genetic propensities related to educational attainment[2] and certain types of SEND. For instance, propensities related to poor educational outcomes, health-related disorders, neurodevelopmental conditions related to having SEND, and mental health/behavioural issues that impact a child's educational attainment.

The key mechanism through which 'Professional' genomics would become a reality in the future would be through screening babies and young children. The current approach to considering whether a diagnostic method is ethically justifiable is to weigh up the cost-effectiveness - whether the harm caused by screening a child is outweighed by the benefit of providing diagnosis and appropriate support/treatment. In the same way that any form of diagnosis could cause potential psychological harm, it is still ethically justified. Therefore, if the evidence suggests genomics could support the diagnosis of certain conditions, that health system could drive demand for this. Some of these conditions may relate to SEND, which would have implications for the education system. There will be strong arguments for 'Professional' genomics to be heavily regulated due to the clinical nature of needs/disabilities being diagnosed.

'Consumer' genomics

'Consumer' genomics relates to the potential use of PGS for assessing non-clinical, non-health-related and less tangible propensities that can impact a child's educational attainment and outcomes, such as testing for IQ or propensity for specific subjects, i.e. maths. The separation of these two markets is not definitive but is useful for separating out the scientific, ethical, and legal differences between the complex variety of products that genomics in education could offer.

The market for genomics in education is currently very small, with only a few USA-based companies offering any kind of genetic screening related to educational attainment, such as genome based/genetic IQ tests (Regalado, 2018). Based on the literature exploring

[2] Educational attainment is used as a cluster term in the academic literature to define a number of educational outcome measures, although there is not necessarily a universally agreed upon meaning – it has been defined as the number of years in education, progression into higher levels of education, high performance in education. It can refer to SEND also, though we have tried to distinguish any outcomes related to SEND.

direct-to-consumer testing (DTCT) in other markets, DCGT companies could take advantage of technological improvements in genomics in education, utilise advertising campaigns to target parents and benefit from growing distrust in the NHS/education system. DCGT is currently offered by a limited number of private companies, such as 23AndMe, which appear to have PGS for educational attainment in mind. However, commercial companies are likely to soon start offering such testing. DCGT companies have already been accused of overselling the accuracy of genetic health tests, and it is important to note that companies' reputation is based on scientific validity.

There is a growing trend for personalised support in healthcare and education, which is tied to the influence of the Silicon Valley-dominated DCGT market (Hogarth & Saukko, 2017). Growing demand for this type of hyper-individualistic diagnostics presents significant issues for a public system like the NHS and the UK education system, particularly in the context of growing distrust in public institutions. There is already evidence to show that genetic screening for health outcomes is causing problems for NHS staff/doctors and the public, so it is likely this will be the same for our education system.

The potential utility of genomics technologies in improving educational outcomes

Genomics could be used to improve educational outcomes by identifying those at risk of poor educational attainment, neurodevelopmental conditions, and/or mental/psychological disorders. One study compares the potential use of PGS to assess educational attainment with the current use of the pupil premium which provides additional funding to schools for each pupil who has been in receipt of free school meals (FSM) in the last six years or has been looked after, adopted or taken into care (Asbury, et al., 2021). The government states that the purpose of the pupil premium is to provide additional funding to schools to help them improve the attainment of disadvantaged children. Effective use of genomic technologies could look like additional resources being allocated to schools or pupils with increased risk of poor educational outcomes and/or neurodevelopmental conditions. Although it is worth stating that the study itself draws mixed conclusions about such a comparison, it does give an idea of what effective use of genomics in education for improving educational outcomes could look like.

There is a complex interplay between environment and genetics in determining educational achievement and attainment. With this in mind, genomic information could be used to assess risk factors and therefore support the design of early intervention to improve educational outcomes for children with ASC, ADHD, dyslexia, and other special educational needs. Children with genetic propensities for developing SEND or even mental health/behavioural issues could be identified in preschool. This could help educators put strategies in place to support them through their educational journey.

The literature highlights the importance of not excluding environmental factors. However, it suggests that genomic technologies could shed light on the complex interplay between genetics and environmental data as they include gene-environment correlations and highlight complex developmental processes. According to some research, Genomics could improve educational outcomes by improving the database and understanding of the science of gene-environment interplay. Genomics, in this sense, could help improve early interventions by further expanding our knowledge of child development (Asbury, et al., 2021).

A recent DfE report on educational psychologists showed that they feel their workload is disproportionately geared towards individual Autism Spectrum Condition (ASC) assessments and other forms of diagnosis over systemic work, which relates to improving the learning environment and identifying the effectiveness of a schools' Special Educational Needs and behaviour support systems. Increasing the ability of staff to focus more time on systemic work has the potential to improve educational outcomes for a greater number of children in need (Department for Education, 2023). Conceptually, genomics in education has the potential to improve the support for those at an elevated risk of developing a SEND by reducing waiting lists for children seeking a diagnosis of SEND or by streamlining diagnostic systems so that education staff can prioritise systemic work and plan interventions proactively rather than reactively.

There is a complex interplay between environmental and genetic factors and education could be the best setting to implement genetic screening. One USA study looks at the impact of performing genetic testing on a cohort of young people whilst teaching them about the science of genetic testing in their biology classes (Gason, et al., 2006). The study showed how young people can have the opportunity to be educated on the science of genomics in a way that limits stigmatisation, promotes agency, and increases visibility as well as support. This also gives young people the opportunity to decline being tested. Although this would not work in the case of genetic screening from pre-school, including genomics within the curriculum nationally could be a way to deal with the problem of stigmatisation and communication of genomics in education whilst also improving educational outcomes.

The literature provides examples of how genomics could be effective in the education space over time. The advancement of AI technologies is already contributing and will continue to contribute to accelerating progress in personalisation and prediction capabilities in genomics. The use of AI has benefits, for example, making genomics technologies more accessible and less costly. Conversely, it brings disadvantages of AI-related discrimination and oversight (The Nuffield Council on Bioethics, 2023).

Key issues and concerns

While there is potential utility to using genomics in educational contexts, there are concerns about maintaining privacy, the prevalence of deterministic views, and how

using genomics in education may impact underprivileged groups, such as those from lower socioeconomic backgrounds and ethnic minority backgrounds. An overarching ethical dilemma raised in the literature is the potential for it to entrench inequalities in the education system and society.

Determinism and Inequalities

The first ethical concern relates to genomics in education leading to 'determinism', that:

- 'heritability' within the context of genomics in education would be misinterpreted as meaning that environmental factors in attainment and outcomes are unnecessary (Cesarini & Visscher, 2017); and,
- 'heritability' relates strictly to a population at one point in time and therefore, ignores, changes across time and space.

In relation to this concern, research has shown the view that genetic data is deterministic is already prominent in the general population. The implications of such a view are groups of children being denied access to education due to misperceptions of how genetics data can be used. Similarly, teachers may label young people based on their genetic information, creating long-term negative impacts. This means that there is a risk of structural issues such as poverty being ignored in favour of putting more effort into genetic issues.

The literature also highlights the connotations that come with any discussion or application of genetics due to the historical use of genetic arguments. Historically, language related to genetics has been used to describe racial and socioeconomic differences for factors such as intelligence whilst ignoring structural elements (Martschenko, et al., 2019). Such a discourse has also been used to resist desegregation, immigration, and to validate socioeconomic and racial inequalities as non-structural issues. This is problematic as it creates the implication that no policy can change genetics, therefore promoting a deterministic view. Even more problematic would be the implication of punitive policies. This is especially concerning given the lack of representation of ethnic minorities in research samples and datasets, meaning in reality that the 'personalised education' would only be useful for specific groups of people.

There is also a concern that the benefits of using genomics in education for personalising education would only accrue to a select group of people and that there would be differing impacts on those from lower socioeconomic backgrounds as well as those from ethnic minority groups. If genomics is commercialised within the education space, those who have higher-risk genetic backgrounds and are living in adverse environments may experience a disadvantage due to the costs associated with accessing genetic tests. This could lead to missed opportunities in monitoring and improving their educational settings, which may, in turn, widen the opportunity gap for children from different socio-economic backgrounds (Genomics England). Conversely, if genomics is made publicly available to all and monitored through governmental bodies, it may mean that those from more

disadvantaged backgrounds are subjected to more surveillance. This group would, therefore, be at more risk of their private information being disclosed (Sabatello, 2018).

Privacy

There are three broad concerns regarding privacy. Firstly, there are concerns about sharing and managing large-scale datasets internationally that respect national boundaries (Birney, et al., 2017). Second, in the education setting (vs. clinical setting), it is much more unclear who has the right to access students' data – for example, in the USA, schools can have access to students' medical data if there is a legitimate interest for them to do so (UK National Screening Committee, 2024). Third, anonymisation of personal genomics data is particularly difficult in comparison to other forms of data. This is especially true as genetic information is shared between relatives and ethnic groups, making it possible to narrow down the identity of a person even in an anonymised database (Lewis, et al., 2020).

Impact on different audiences

The advancement of genomics in education can impact children and young people, parents, and educators in various ways. Whilst some of the implications are clear for parents and young people (such as being able to diagnose certain disorders quicker), the social and emotional implications have not been explored. Similarly, beyond the ethical implications already discussed, very limited research discusses the potential adverse effects of using genomics within the education space and the impact on different audiences is not widely understood given the stage of development. However, a summary of what the literature hypothesises is discussed below.

Children could potentially be diagnosed earlier in life for disorders such as Autism, ADHD and dyslexia (Johnson, et al., 2013). Similarly, health conditions that may impact their educational attainment could be discovered sooner. Both of these mean that young people could have the right support and preventative measures in place. Having this knowledge upfront and putting support measures in place means that children and their parents have an increased sense of agency (Finlay, 2017).

On the other hand, it is important to note that the social, emotional, and cognitive impact of receiving a diagnosis has not been thoroughly explored in previous research. Furthermore, the main benefit of using genetic testing to diagnose any form of SEND is reliant on the education system's ability to provide appropriate support. Additionally, it would be very important that, if diagnosis through genetic testing for SEND was widespread, ableist discrimination and the pigeon-holing of children based on such diagnosis did not occur. Children with SEND already face higher rates of bullying and disenfranchisement with the education system. Genomics in education could potentially add to these issues if not managed/organised responsibly.

Similar to young people, parents would have an increased sense of agency by knowing certain health issues and disorders so they can take action to mitigate the impact. If schools would not be able to adopt educational models based on genomics data, however, this could cause frustration for parents who may feel the information they have is not being taken seriously. There is also a risk that parents would not understand the limitations of the information they are receiving.

Teachers would need significant amounts of information and training on how to work more effectively with children who have specific syndromes and health-related issues and gain a deeper understanding of the effectiveness of different educational methods with various groups. On the other hand, teachers wouldn't fully understand the gene-environment interaction and would therefore require professional development to utilise genetic findings effectively (Morris, et al., 2024).

Annex B: Future Pathways

Likely developments in the field of genomics relevant to education

Genomics in education is the study of the human genome applied to the study of educational attainment and education-related issues. These future pathways have been developed to follow two potential strands of the application of genomic technologies to educational settings. These two pathways are not mutually exclusive but rather attempt to lay out, from the evidence review (Annex A) and stakeholder workshop, two distinct, potential and realistic futures for genomics in education within the next 10 years, assuming the science progresses as it has been so far.

Pathway 1: Further Developments in Professional Genomics

The scientific research could improve to a point where it becomes feasible to use genomics (i.e. Polygenic Risk Scores) to assess the genetic risk more accurately, at an individual level, of developing conditions categorised as Special Educational Needs and Disability (SEND). These include Autistic Spectrum Disorder (ASC), *Attention Deficit Hyperactivity Disorder (*ADHD) and dyslexia. It is important to note that the 'Professional' pathway includes different applications of genomic technologies. As prior discussed, the use of PGS and other screening mechanisms can be used to assess risk of certain conditions like ASC from an early age. However, this approach to other conditions may not be as useful for conditions such as for ADHD which research suggests may be more subject to environmental factors over a person's youth to adulthood. In such a case, some literature indicates the utility of a sub-field of genomics (pharmacogenomics) to assess genetic propensities for certain medications.

This has been termed as 'Professional' genomics because it relates only to conditions that require clinical medical diagnosis and remediation. This assessment of genetic propensity is complex and applies to neurodevelopmental conditions differently, as differing genetic and environmental factors interact. The influence of environmental factors is highly significant for predicting the development of neurodevelopmental conditions.

Genetic research is already taking place across a number of health/NHS settings, including the D-CYPHR study, UK Biobank, Our Future Health and the Generation Study (also known as whole genome sequencing programme). It is likely that one of these existing programmes will develop an effective mechanism for developing and rolling out the technology. Genetic tests for SEND would allow for earlier identification and the potential for targeted support and intervention for children at risk of developing such conditions. This may offer a more systematic way for identifying risk of SEND, so less

children are diagnosed later in life. There is broad agreement that early intervention and early support is better than later stage diagnosis.

Risks

The use of ‘Professional’ genomics in educational settings poses several risks, however:

- Unless all stakeholders involved are well educated around the difference between a diagnosis and a risk score, and thus have sufficient understanding of complex genetic information, they may view genetics and genetic risk scores as deterministic.
- More broadly, but linked to this, genetic screening for such conditions could lead to an oversimplification of complex behavioural conditions such as ASC, ADHD, and dyslexia, potentially ignoring the significant role of a child’s environment in their development.
- Children identified at risk of developing SEND may be treated differently (i.e. labelled, discriminated against, or stigmatised), and this may affect how people respond to them throughout their lives. For example, the information may influence teachers’ behaviour towards the child, reinforcing behaviours and leading to a form of ‘self-fulfilling prophecy’.
- This use of genomics would most likely require screening non-consenting individuals and storing their data up until they withdraw consent. Without consent as a basis, the ethical and corresponding legal justification for using genomics in this manner requires an assessment of harm caused to the child over the potential benefit of diagnosis and consequent support. Public distrust in the moral/legal justification for screening children could also produce many negative outcomes.
- Current over-representation of people of European descent could cause issues where people of African and Asian descent could receive less reliable or incorrect diagnosis of risk. Especially amongst ethnic groups where there is higher distrust in the education system, this could cause further disengagement. In general, the under-representation of non-European groups causes many issues for ‘Professional’ genomics.

Considerations

The use of ‘Professional’ genomics in educational settings poses several considerations:

- Larger data sampling and significant research must be undertaken for this use of genomics to be developed with sufficient accuracy and reliability. Additionally, there would need to be sufficient monitoring and evaluation of how any genomics programme functions, and what impact this has on those involved, especially children.

- The use of genomics to identify a higher risk of developing SEND is different to a diagnosis of a physical condition like asthma. Accordingly, if genomics were part of the system for SEND identification and support, there would need to be further consideration of how to build literacy and understanding of genomic results. For example, genomic counselling being offered in schools.
- In addition to this, considerations should be made on upskilling of education staff to better understand the use of genomics in education. This may promote a better understanding of how teachers will want to use genetic information, alongside other information available about children in educational settings.
- Increased identification of SEND through genomics would increase expectations around the provision of services and support for such children.
- Using genetics in educational settings, particularly around SEND where it could be feasible (as opposed to educational attainment genetic traits) could reinforce the societal position of trying to 'fix' children with such conditions, potentially limiting ambition and outcomes, as opposed to celebrating and accommodating for differences between children.
- For this use of genomics to be successful, it would require the collection and storage of genomic data from non-consenting individuals. This data is hard to anonymise and would be highly valuable for private companies. Accordingly, many considerations on how this data is collected, stored, protected, used and destroyed in accordance with thorough data protection legislation is highly important.

Pathway 2: Further Developments in 'Consumer' Genomics

Unregulated Direct-to-Consumer Testing Companies (DTCT) are likely to start marketing genetic tests for 'Consumer' genomics. This will relate to the use of genetic tests (i.e. polygenic risk scores) for assessing non-clinical, non-health-related and less tangible propensities that can impact a child's educational attainment and outcomes, such as cognitive ability, mathematical/logistical capabilities, reading ability and even IQ. Although, it is important to state that DTCT companies potentially could also offer screening related to 'Professional' genomics, as some DTCT companies do already offer some health-related genetic screenings for conditions such as dementia. There is already an existing precedent for public leadership in the space of 'Professional' genomics as well as the nature of 'Professional' genomics, for instance 23andMe offers a test for health-related conditions but have a legal disclaimer that no government entity or UK/EU based regulatory bodies approve of the clinical validity of such tests. It is likely that DTCT companies will focus more heavily on 'Consumer' genomics because they will not have to offer such legal disclaimers on clinical issues, consequently, facing less legal/regulatory pressures. Arguably, 'Consumer' genomics relates to more marketable measures of educational attainment that will cause the least backlash from the general public. These

tests will cover a variety of measures for educational attainment and cognitive ability, some of which do have less scientific backing and research to support their validity. The data used to develop these tests is likely to be biased, overrepresenting population groups of European descent and underrepresenting other population groups and ethnicities.

DCGT companies could utilise powerful advertising campaigns to target parents and benefit from growing public distrust and frustrations in the NHS and education system. However, such companies are also aware of the ramifications to their brand if there is growing consensus from consumers that their tests are unreliable. It is currently difficult to predict what is most realistic, given the DCGT market for health-related conditions is still very young, having only just started in the past few years.

Risks

The use of 'Consumer' genomics in educational settings poses several risks, however:

- DCGT companies will likely market tests regardless of their accuracy or reliability. These companies have already been accused of overselling the accuracy of genetic health tests.
- Parents could start to buy-in to these tests on mass, without a full understanding around their limitations. Coupled with this, unless genetic literacy within the population substantially improves, parents could put their faith in inaccurate *risk scores* (as opposed to diagnoses).
- Given the costs associated with these tests, only those who can afford them will buy into them, meaning that they will not be equally accessible to all.
- Given these tests would be marketed as 'scientific', parents or caregivers could then bring such test results to teachers/education staff as justification for their child to, for instance, be moved into a higher-performing set, question the validity of their child's grades on exams, or demand further support for their child in certain topics.
- Without guidance/policy on how schools should react to this, this could cause a chaotic response across the UK, additional strain on resources, and further distrust between parents and teachers/education system.
- Given that these are private companies, there are many questions over how these companies protect highly sensitive genetic data. For instance, what happens to people's genetic data if a company folds, how should DTCT companies communicate their terms of service and privacy policies so that consumers understand what they are signing up for, and what ramifications are there when data is stolen or leaked.
- More generally, there are concerns around DTCT companies exploiting information for profit. Genetic data is very difficult to fully anonymise, is life-long,

and highly valuable. For instance, there has been a recent large data breach scandal with 23andMe, where hackers stole data from specific ethnic minority groups. If something similar happens with a much larger sample and relating to people's cognitive abilities, that could cause many issues.

Considerations

The use of 'Consumer' genomics in educational settings poses several considerations:

- Larger data sampling and significant research must be undertaken for this use of genomics to be developed with sufficient accuracy and reliability. It is not currently clear how DCGT companies could collect such large sample sizes via the market alone, as there is not sufficient demand in the general public for such tests. Without regulation, this could mean that these companies offer unreliable and unsubstantiated products to consumers.
- Currently, DCGT companies can market genomics products without necessarily making parents, young people, or educators aware of their limitations, meaning there may be room to inflate what can realistically be concluded from genomics data. With this context, regulators may need to consider how to limit the claims of these companies and what ramifications exist if they break such limits.
- Similar to pathway one, considerations should be made on how to improve genomics literacy and understanding amongst the public and education staff.
- Consideration should be made of what guidance is given to schools, teachers, and educators in response to a parent coming into school with such test results. Although, this would not necessarily have to be policy, some form of national guidance across all educational settings would be most appropriate so that an unequal and chaotic response does not develop across the UK.
- Given DTCT companies would be providing these tests for a cost, those which have higher-risk genetic backgrounds and are living in adverse environments may experience a disadvantage due to the costs associated with accessing genetic tests. This would be due to missed opportunities in monitoring and improving their educational settings, which may, in turn, widen the opportunity gap for children from different socio-economic backgrounds and reproducing inequalities.
- Considerations should be made of how the private DCGT market will interact with government/public institutions. There are tipping points for how much power DCGT companies develop over the next 20 years in this space. If they corner the market, this could lead to many issues for government down the line which they cannot turn back the clock on. Although alternatively, there are positive examples of private-public partnerships producing major advancements in the field, such as with the UK Biobank, it is possible that the future of genomics will involve more collaboration between public and private actors.

- There are also constraints around the capacity within the education system to respond to the demand in assessments for SEND. Increased screening could impact on the already limited capacity.

Annex C: Consumer Market Research

Purpose: this note sets out a summary of the commercial genomics market based on a desk review conducted as part of a futures project on the implications of genomic technologies in education.

Market Overview

The direct-to-consumer (DTC) genetic testing market is an expanding field that allows consumers to access information about their genetics without going through a healthcare provider. This market includes a range of services, from ancestry and heritage discovery to personalised healthcare and wellness advice based on a person's genetic makeup. Key players in this market include 23andMe, AncestryDNA, MyHeritage, Helix, Living DNA, and others. These companies offer a variety of services, from ancestry tracing to health risk assessments.

In terms of market size, according to a report by Credence Research, the global DTC genetic testing market was valued at around $1.24 billion in 2020 and is expected to grow at a compound annual growth rate (CAGR) of approximately 14% from 2021 to 2028. The North American region dominated the market in 2020, due to the presence of key players and high consumer awareness regarding genetic testing.

Several factors, including the increasing prevalence of genetic diseases, the rising awareness and accessibility of DTC genetic testing, advancements in genomics technology, and the increasing interest in personalised medicine drive this market's growth. However, ethical and privacy concerns regarding genetic data, along with the potential for misinterpretation of genetic information by consumers, pose challenges to the market's growth.

The growth in the market for genomic testing for ancestry and health purposes

Genomics, for ancestry/health purposes, forms a significant part of the DTC genetic testing market. Companies like 23andMe and AncestryDNA have popularised the use of genetic testing for tracing lineage and discovering heritage, and this segment of the market has seen substantial growth. The substantial growth in the use of DTC genetic testing for tracing lineage and discovering heritage really started to pick up around the mid-2010s. This growth was driven by several factors, including advancements in genomics technology, decreases in the cost of genetic sequencing, and increased consumer interest in personal genetics and ancestry. For example, AncestryDNA launched its genetic testing service in 2012, and by 2017, it had already tested over 5 million people. 23andMe, which launched its genetic testing service in 2007, saw a similar growth trajectory. It's worth noting that growth rates can vary by region and are

influenced by factors such as regulatory environments and cultural attitudes towards genetic testing. In addition, there's a growing interest in using genetic testing for health purposes, such as identifying genetic risks for certain diseases and planning personalised healthcare based on genetic data.

Marketing

Whilst there is not a significant number of products directly marketed relating to educational attainment, many of the tests provide information that is education-related. Common marketing messages include:

Utility: Marketing messages often highlight the practical benefits of genetic testing. This includes the ability to uncover one's ancestry, discover genetic relatives, and gain insights into personal health risks and traits. Companies emphasise how understanding one's genetics can inform lifestyle choices, health decisions, and provide a greater understanding of one's identity.

Validity: To build consumer trust, companies often emphasise the scientific rigor and accuracy of their tests. They highlight the use of advanced genomics technology, large reference databases for ancestry comparison, and their compliance with relevant regulations. They also often mention partnerships or collaborations with research institutions to validate their testing methods and findings.

Educating customers about Genomics: The educational aspect is a significant part of the marketing strategy for these companies. They stress how their products can help consumers learn more about genetics in a personal and engaging way. This goes from understanding the basics of DNA, to learning about genetic heritage, to interpreting what genetic variations may mean for personal health.

For instance, 23andMe often emphasises how their product can be used as an educational tool for understanding genetics and health. They highlight their online education module, which breaks down complex genetic information into understandable insights. On the other hand, AncestryDNA focuses more on the educational aspect of learning about one's family history and cultural heritage. They often feature customer stories about discovering unknown ancestors or cultures in their marketing campaigns. To sum up, the general marketing message of DTC genetic testing companies is that their products are not only useful and scientifically valid but also serve as an engaging and personal way to educate oneself about genetics, heritage, and health.

Customers

Common traits of customers in the DTC market:

Age Group: The customer base for DTC genetic testing tends to skew towards the younger side, particularly among individuals in the age bracket of 20-45 years. This could be attributed to this group's increased propensity for adopting new technologies and their higher interest in personalized health and wellness.

Gender: While both genders are represented in the customer base, some reports suggest a slight female predominance. This could be due to women's generally higher engagement with healthcare and wellness initiatives.

Geographic Distribution: North America, particularly the United States, has the largest customer base in the global DTC genetic testing market. This is due to a combination of factors like the presence of major market players, high awareness levels about genetic testing, and relatively higher disposable income.

Education Level: Individuals with higher education levels are more likely to use DTC genetic testing services. This could be due to their increased understanding of the potential benefits and implications of genetic testing.

Ancestry vs. Health: According to a study published in the Journal of Personalized Medicine, more consumers use genetic testing for ancestry-related purposes than for health-related ones. However, the interest in health-related genetic testing is growing, particularly among individuals with a family history of genetic diseases.

Privacy Concerns: Privacy is a significant factor for customers in this market. A study in the journal Nature found that people who had privacy concerns were less likely to have used or intended to use DTC genetic testing.

Companies

23andMe, one of the most prominent companies in this field, reported more than 12 million customers as of 2021. Their revenue in 2020 was approximately $305 million. Their services centre around two main testing kits: The Health + Ancestry Service and the Ancestry + Traits Service. The former provides insights into a user's health predispositions, carrier status for certain conditions, wellness, and ancestry. The latter focuses on ancestry and traits analysis.

AncestryDNA, another major player, reported a customer base of 18 million people in 2020. Their primary product is the AncestryDNA kit, which provides insights into the user's ethnic mix and the regions of the world their ancestors come from. They also offer an AncestryHealth kit that provides health insights based on genetics. The company's revenue in 2020 was around $1 billion.

MyHeritage, with a user base of 4.4 million as of 2020, offers a DNA test that provides ethnicity estimates and helps users find relatives around the world based on shared DNA. They also offer a Health kit, which includes all the benefits of the DNA kit plus a

comprehensive health report. The company's 2020 revenue was approximately $141 million.

Annex D1 gives a brief summary of the key players, and the case studies below illustrate the approach of two prominent companies.

Case Study 1: 23andMe's Educational Approach to Genetic Testing

23andMe is one of the pioneers in the DTC genetic testing market. Over the years, they have differentiated themselves with a strong emphasis on the educational aspect of genetic testing, helping their customers understand the implications of their genetic health risk reports. One of their unique offerings is an online education module that breaks down complex genetic information into understandable and actionable insights. This module provides a comprehensive guide on how to interpret genetic health risk reports, what the results mean, and how they can be used to make informed decisions about personal health and wellness. In terms of marketing, 23andMe often highlights the educational benefits of their products. They utilize various channels, including social media, blogs, and direct mail, to communicate how customers can learn about their ancestry, traits, and health predispositions through their genetic testing kits. This strategy not only empowers consumers with knowledge about their genetics but also positions 23andMe as a reliable and trustworthy provider in the DTC genetic testing market.

However, 23andMe faced a data breach which led to a growing distrust among its consumers and the public in general. Such breaches are a massive violation of privacy, and it can have long-lasting impacts on an individual's life. In the case of 23andMe, the data breach exposed the genetic information of their customers, which can potentially be used for nefarious purposes. The incident has had a significant impact on public perception. Firstly, it has raised concerns about the safety and security measures of DCGT companies. Consumers are now much more cautious and apprehensive about sharing their sensitive data. This distrust, in turn, hampers the growth of the DCGT market as people are less willing to use these services due to fear of their data being misused or exposed. Secondly, it has led to a demand for stricter regulations and better security measures for DCGT companies. Many people are calling for better oversight of these companies to prevent such incidents in the future. This has also put pressure on other companies in the DCGT market to demonstrate that they have robust security measures in place to protect customer data. Lastly, the incident has also had a negative impact on 23andMe's reputation. The company has faced backlash from customers and the public, which can potentially affect its customer base and sales. In the competitive DCGT market, maintaining customer trust is crucial for success. If a company fails to do so, it can have severe repercussions on its market position and revenues, as well as the future of genomic testing.

Case Study 2: AncestryDNA's Heritage Education Approach

AncestryDNA, another dominant player in the DTC genetic testing market, takes a slightly different approach. Rather than focusing solely on health implications, they place a strong emphasis on the educational aspect of learning about one's family history and cultural heritage. Their primary product, the AncestryDNA kit, offers detailed insights into a user's ethnic mix and geographical origins. They have a vast database, which allows them to connect customers to regions around the world where their ancestors hailed from. This has led to many customers discovering previously unknown ancestors, roots, or cultural backgrounds. AncestryDNA's marketing campaigns often revolve around these success stories, highlighting how their products have helped customers delve deeper into their family histories and cultural heritage. They use a variety of marketing channels, including television commercials, online ads, and social media, to share customer testimonials and emphasize the educational value of their service.

Annex D1: Overview of key market players

23andMe

23andMe is a leading player in the DTC genetic testing market. Founded in 2006, the company aims to help people access, understand, and benefit from the human genome. They have pioneered the development of DNA analysis services for consumers.

- Their Health + Ancestry Service provides insights into a user's health predispositions, carrier status for certain conditions (their website does not mention SEND), wellness, and ancestry.
- The Ancestry + Traits Service focuses on providing detailed reports on the user's ancestry and personal traits, offering insights into their heritage and genetic makeup.

AncestryDNA

AncestryDNA, a subsidiary of Ancestry.com, is a global leader in digital family history services. They leverage advanced DNA science to deliver detailed genetic reports.

- Their primary product, the AncestryDNA kit, provides insights into the user's ethnic mix and geographical origins of their ancestors.
- AncestryHealth, another product, offers health insights based on genetics, providing personalised health plans, and identifying potential health risks.

MyHeritage

MyHeritage is a platform that combines genealogy with genetic testing. Founded in 2003, it allows users to create family trees, upload and browse through photos, and search billions of global historical records.

- The MyHeritage DNA test provides ethnicity estimates and helps users find relatives around the world based on shared DNA.
- MyHeritage Health offers a comprehensive health report in addition to the benefits of the DNA kit.

Helix

Helix is a genomics company that aims to make DNA learning accessible and actionable for everyone. They provide a marketplace of DNA-powered products addressing a wide range of categories like health, ancestry, entertainment, family, fitness, and nutrition.

- The Geno 2.0 Next Generation Ancestry DNA Test, powered by Helix and developed by National Geographic, provides insights into the regional ancestry and the migration patterns of the user's ancestors.
- The Mayo Clinic Gene Guide, offers insights into how genetics can influence health, providing actionable recommendations for lifestyle changes and preventive measures.

Living DNA

Living DNA is a UK-based genomics firm that provides DNA testing for ancestry and wellness. Their tests are designed to deliver detailed insights into the user's ancestry and genetic health.

- The Living DNA Wellbeing Kit is designed to help users understand their body's response to different lifestyle choices, enabling them to personalise their health and wellness strategies.
- The Living DNA Ancestry Kit provides a detailed view of the user's lineage, tracing maternal and paternal lines separately to give a comprehensive view of the user's family history.

Bibliography

Allegrini, A. G. et al., 2019. Genomic prediction of cognitive traits in childhood and adolescence. *Molecular Psychiatry 2019 24:6,* 4, 24(6), pp. 819-827.

Anderson, J. S. et al., 2019. Polygenic risk scoring and prediction of mental health outcomes. *Current Opinion in Psychology,* 6, Volume 27, pp. 77-81.

Asbury, K., McBride, T. & Rimfeld, K., 2021. *Genetics and early intervention: Exploring ethical and policy questions,* s.l.: Early Intervention Foundation.

Balogh, L., Pulay, A. J. & Réthelyi, J. M., 2022. Genetics in the ADHD Clinic: How Can Genetic Testing Support the Current Clinical Practice?. *Frontiers in Psychology,* 3.Volume 13.

Birney, E., Vamathevan, J. & Goodhand, P., 2017. Genomics in healthcare: GA4GH looks to 2022. *BioRxiv.*

Cesarini, D. & Visscher, P. M., 2017. Genetics and educational attainment. *npj Science of Learning 2017 2:1,* 2, 2(1), pp. 1-7.

Department for Education, 2023. *Educational psychology services : workforce insights and school perspectives on impact.,* s.l.: s.n.

Finlay, T., 2017. Testing the NHS: the tensions between personalized and collective medicine produced by personal genomics in the UK. *New Genetics and Society,* 7, 36(3), pp. 227-249.

Gabbay-Dizdar, N. et al., 2022. Early diagnosis of autism in the community is associated with marked improvement in social symptoms within 1–2 years. *Autism,* 8, 26(6), pp. 1353-1363.

Gason, A. A., Delatycki, M. B., Metcalfe, S. A. & Aitken, M. A., 2006. *It's 'back to school' for genetic screening,* s.l.: Nature Publishing Group.

Genomics England, n.d. *Newborn Genomes Programme.* [Online]
Available at: https://www.genomicsengland.co.uk/initiatives/newborns

Gialluisi, A. et al., 2020. Genome-wide association study reveals new insights into the heritability and genetic correlates of developmental dyslexia. *Molecular Psychiatry 2020 26:7,* 10, 26(7), pp. 3004-3017.

Government Office for Science, 2022. *Genomics Beyond Health,* s.l.: GOV.UK.

Grove, J. et al., 2019. Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics 2019 51:3,* 2, 51(3), pp. 431-444.

HM Government, 2022. *SEND Review: Right support Right place Right time A guide for children and young people to the special educational needs and disabilities (SEND) and alternative provision green paper,* s.l.: s.n.

Hogarth, S. & Saukko, P., 2017. New Genetics and Society Critical Studies of Contemporary Biosciences A market in the making: the past, present and future of direct-to-consumer genomics.

Jansen, P. R. et al., 2018. Polygenic scores for schizophrenia and educational attainment are associated with behavioural problems in early childhood in the general population. *Journal of Child Psychology and Psychiatry,* 1, 59(1), pp. 39-47.

Johnson, N. L., Giarelli, E., Lewis, C. & Rice, C. E., 2013. Genomics and Autism Spectrum Disorder. *Journal of Nursing Scholarship,* 3, 45(1), pp. 69-78.

Lee, J. J. et al., 2018. Gene discovery and polygenic prediction from a 1.1-million-person GWAS of educational attainment. *Nature genetics,* 8, 50(8), p. 1112.

Lewis, C. et al., 2020. Parents' motivations, concerns and understanding of genome sequencing: a qualitative interview study. *European Journal of Human Genetics,* Volume 28, pp. 874-884.

Martschenko, D., Trejo, S. & Domingue, B. W., 2019. Genetics and Education: Recent Developments in the Context of an Ugly History and an Uncertain Future. *https://doi.org/10.1177/2332858418810516,* 2, 5(1), pp. 1-15.

Morris, T. T. et al., 2024. Implications of the genomic revolution for education research and policy. *British Educational Research Journal,* 6, 50(3), pp. 923-943.

NHS England, n.d. *NHS Genomic Medicine Service.* [Online]
Available at: https://www.england.nhs.uk/genomics/nhs-genomic-med-service/

Rabinowitz, J. A. et al., 2019. Associations between an educational attainment polygenic score with educational attainment in an African American sample. *Genes, Brain and Behavior,* 6, 18(5), p. e12558.

Regalado, A., 2018. DNA tests for IQ are coming, but it might not be smart to take one | MIT Technology Review. 4.

Sabatello, M., 2018. A Genomically Informed Education System? Challenges for Behavioral Genetics. *Journal of Law, Medicine & Ethics,* 3, 46(1), pp. 130-144.

Selzam, S. et al., 2016. Predicting educational achievement from DNA. *Molecular Psychiatry 2017 22:2,* 7, 22(2), pp. 267-272.

The Nuffield Council on Bioethics, 2023. *The application of AI to genomics raises 'urgent' ethical questions, early findings suggest.* [Online]

Available at: https://www.nuffieldbioethics.org/news/the-application-of-ai-to-genomics-raises-urgent-ethical-questions-early-findings-suggest

UK BioBank, n.d. *Why have we sequenced half a million genomes?.* [Online]
Available at: https://www.ukbiobank.ac.uk/explore-your-participation/understanding-genetics/why-have-we-sequenced-half-a-million-genomes#

UK National Screening Committee, 2024. *UK National Screening Committee: screening in healthcare - Principles of screening - Guidance - GOV.UK.* [Online]
Available at: https://www.gov.uk/guidance/principles-of-population-screening/principles-of-screening

Weiner, D. J. et al., 2017. Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. *Nature Genetics 2017 49:7,* 5, 49(7), pp. 978-985.

Zhou, M., Zhang, Y.-M. & Li, T., 2023. Knowledge, attitudes and experiences of genetic testing for autism spectrum disorders among caregivers, patients, and health providers: A systematic review. *World Journal of Psychiatry,* 5, 13(5), p. 247.

HM Government

Reference: RR1443

ISBN: 978-1-83870-584-8

For any enquiries regarding this publication, contact www.education.gov.uk/contactus.

This document is available for download at www.gov.uk/government/publications.