

CLOSING THE GAP IN PATIENT DATA FOR BLACK AND SOUTH ASIAN COMMUNITIES:

A guide for NHS leaders and policy professionals

This guide provides an introduction to the inequalities in patient health data for Black and South Asian communities. It shares practical solutions, co-designed by healthcare workers and communities, to support people who work for the NHS to play their role in reducing health inequalities by improving their data practices.



ACKNOWLEDGEMENTS

This resource was co-designed with members of the community and healthcare workers from Black and South Asian backgrounds.

We want to acknowledge our co-authors:

Agnes Agyepong Sabrina Phillips

Harjinder Singh Stephen Quashie

Mervyn Kennedy-Macfoy Valerie C. White

Neheed Akhtar Theshnee Govender

Olubunmi Adekoya

Creary, N, Quashie, S. Kennedy-Macfoy, M, Singh, H, Adekoya, O., Phillips, S., Akhtar, N., Govender, T., White V C., Agyepong, A. (2022), Closing the gap in patient data for Black and South Asian communities: A guide for NHS leaders and policy professionals: Liberating Knowledge.

https://understandingpatientdata.org.uk/insights-black-south-asian-people-patient-data

I'D LIKE TO KNOW...

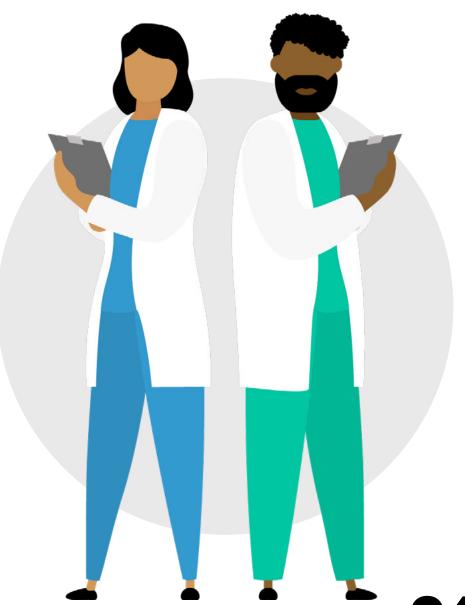
What this guide is about	04
What is patient data?	05
How patient data is used	<u>07</u>
Why we need to reduce gaps in patient data for Black and South Asian communities	08
The barriers to collecting and making best use of data	09
How to reduce the risk of causing harm	12
How to earn trust and build confidence	16
What I can do to improve data practices?	<u>17</u>
How to support patients to decide how their data is collected and used	21
Where I can find more information about patient health data?	24
More about some of the terms and concepts covered in this guide	25

WHAT THIS GUIDE IS ABOUT

This guide has been created to raise awareness of the negative consequences that gaps in patient data have for the health outcomes for Black and South Asian communities. It explores the reasons put forward by healthcare workers and community members to explain why inequalities in data collection and use exist.

This resource provides an overview of how data is collected, used and stored. It shares practical solutions for healthcare workers to support communities to make informed decisions about how their data is used. It also highlights how senior leaders and key decision makers can use their power and influence to support the cultural and structural changes required to improve data practices within healthcare.

We hope this guide provides a useful starting point for the healthcare sector and communities to come together to improve data practices that affect patients from Black and South Asian backgrounds, to provide better care experiences and improve health outcomes for us all.



WHAT IS PATIENT DATA?

Patient data collected by healthcare services is information about a person's physical and mental health. This information could be about their past medical history or current health status. It includes information needed to provide the services a patient has requested, or to access the care they need. However, sometimes it is less clear why other forms of patient data are collected.



We have provided an example of the types of personal data that may be collected on a patient:

- Clinical data (e.g. health tests, weight, blood pressure, etc.)
- Date of Birth
- Ethnicity
- Address
- Notes written by the healthcare professional

- Sex
- Religion
- Sexual orientation
- Relationship status
- Education
- Health app data
- Patient surveys
- Records of complaints



This data is collected through a number of ways - for example through having conversations with a receptionist in a hospital, during a consultation with a GP, interactions with a nurse during a home visit, patient surveys, referrals received from organisations that work in partnership with the NHS (e.g. social care) etc. To make a positive impact on services and the health of the population patient health information needs to be recorded accurately, analysed and used consistently across teams and organisations. This makes it easier for the NHS to identify good practice and where services are working well, as well as areas for improvement.

To protect patients' privacy the data is de-identified, the data is de-identified which removes personal information, assigns general identifiers or combines information from a number of patients to create a broader group or category. The process of de-identifying the information means it is difficult to identify individual patients. These large datasets enable health services to explore trends in patients' access, experience and outcomes in services. However, in rare circumstances when data has been de-identified it still may be possible to re-identify someone if they have unique characteristics or social identities that are under-represented. To manage this, data on samples smaller than 10 patients would not be published (for further explanation see pages 15 to 18 of What happens to health data).¹

HOW PATIENT DATA IS USED

Most people working in healthcare are aware of the range of data that is collected by the NHS. They also have a good knowledge of the reasons for collecting the data they use. However, some are less clear about how the data they are not responsible for is used and managed by the NHS. This in turn makes it difficult to explain to patients why their data is needed, especially when the information does not appear to be directly related to the quality of care they receive (e.g. religion, financial situation, relationship status etc.).²

When patient data is recorded accurately, analysed effectively and the information that is generated is acted upon, it has potential to transform patients' experiences of health and services, especially for groups who experience the poorest health outcomes. It makes it easier for organisations to make informed decisions about the services they need to invest in, which could lead to improvements in health for everyone.

Good quality data:



Increases the likelihood that a patient's individual care needs will be met



Enables the NHS to plan services that are designed to meet the needs of the populations they serve



Improves methods and efficiency in diagnosing illnesses



Helps services to provide good quality and safe healthcare



Supports research into new and improved approaches to treat and prevent illness or disease



Can be used to develop better policies which influence our access to healthcare

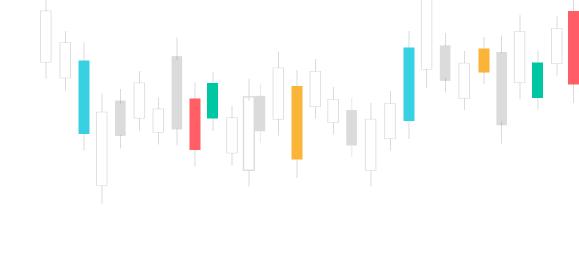


Helps healthcare services to understand how factors outside of the health service (e.g. housing, debt, employment etc.) may impact the health outcomes of their patients

WHY WE NEED TO REDUCE GAPS IN PATIENT DATA

for Black and South Asian communities

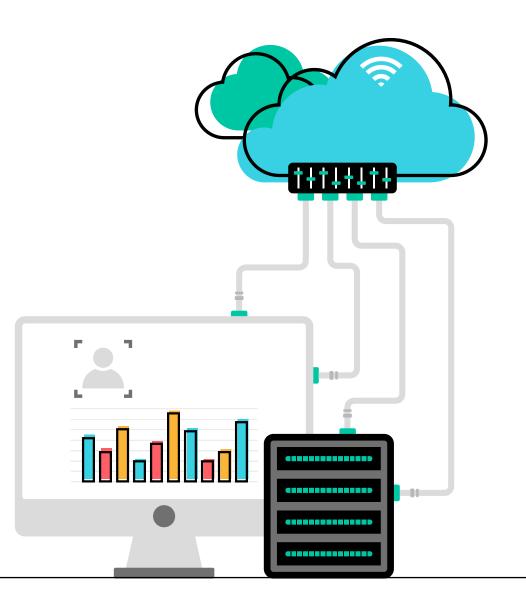
Data collection practices across health services are variable, but there is a worrying trend that ethnicity data collected for Black and South Asian communities has more gaps and inconsistencies than other ethnic groups in the UK.³ Poor data practice is an example of how structural inequalities are encoded into NHS datasets. Data that is missing or poorly recorded can make it difficult for health services to understand and respond to the needs of their patients.⁴ This might mean they receive a poorer experience, develop long term conditions, their conditions may not be managed effectively which could lead to people dying prematurely. Therefore health inequalities are reinforced and wider social injustices.^{5;6}



The data that is available shows that people from Black and South Asian backgrounds often face greater barriers to accessing the care they need, and are more likely to have poorer experiences of services. This is often due to services not being culturally relevant and not designed to meet their needs. They are also more likely to experience racism and other forms of discrimination (e.g. islamophobia, sexism etc.) within services, but also in other aspects of their lives which may make it difficult for them to stay well or get the support they need and this often leads to poorer health experiences. However, due to missing data most services do not fully understand the extent of the inequalities these communities face. Collecting patient data is a crucial step to ensure that they have equal opportunity to live in good health and have access to quality and safe services.

THE BARRIERS TO COLLECTING AND MAKING BEST USE OF DATA

Patient datasets need to be as complete as possible for the NHS to have a true picture of the needs of the communities they serve. Our research and the work of other organisations has identified a range of barriers to data collection.



THE BARRIERS TO COLLECTING AND MAKING BEST USE OF DATA

Barriers expressed by healthcare workers



Despite high levels of awareness about data being collected to deliver services to a patient, there is lack of awareness about how it is also used in research or service planning. This makes it difficult to explain to patients why certain information is requested



Concerns about offending patients when asking for sensitive information (e.g. sexual orientation)



Fear that they may be perceived by patients as being racist if they ask them to provide information about their ethnicity



Not having access to technology in their working environments which would make it easier to record information



When time is limited, many feel it is more important to prioritise providing care to a patient rather than completing forms



A lack of commitment from senior leadership to focus on addressing health inequalities for racially minoritised groups



Concerns that sharing sensitive information about a patient with another department might increase the patient's risk of being stigmatised or experiencing discrimination



Systems and processes do not enable people to analyse data by ethnicity and other <u>protected characteristics</u>⁸



Not clear how the data they collect leads to service improvements or reducing health inequalities



Databases used by some NHS organisations are difficult to use and recording information can be time consuming



Lack of representation of healthcare workers from Black, Asian and other minoritised backgrounds in senior leadership roles²

THE BARRIERS TO COLLECTING AND MAKING BEST USE OF DATA

Barriers expressed by people from Black and South Asian backgrounds



Limited options on equalities monitoring forms do not enable people to accurately self-identify. (e.g. someone may wish to identify as Black African and Zambian; Sikhs also see themselves as a racial group not just a religious group)



Less informed about how their data contributes to improving health services and outcomes for their community and wider society



Most people have not had a conversation with a healthcare worker about why their information is collected and how it will be used



Low levels of trust in the NHS, concerns about data security and their information being exploited by private companies



Concerns that the NHS is not using data to improve the health outcomes for their communities



Negative experiences within the healthcare system make some people fearful that providing information may lead to racial discrimination



Concerns that their data may be weaponised whereby treatment may be forced upon them or withheld



Fears that their data may be shared with external agencies which may put them or their family at risk. (e.g. sharing information with the Police or the Home Office which could lead to harassment and or cause complications for their immigration status)



Some need support to complete forms due to language barriers, lower levels of literacy and learning disabilities



The relationship between communities and the NHS is seen as extractive, when the NHS does not provide feedback about how their data or suggestions have been used to improve services

HOW TO REDUCE THE RISK OF CAUSING HARM

Data is not neutral and contributes to the systemic injustices we see in healthcare and wider society. It plays an important role in what we know about human existence. It shapes the judgements and decisions we make in our daily lives, influences our behaviours and affects how we navigate the world. In the context of healthcare, data shapes who is seen and who is invisible, whose needs are met, overlooked or ignored

(Creary et al, 2022)

Systemic racism is a key driver in creating and sustaining health inequalities. When patient data is used well it can help us to gain a better understanding of patterns of inequality and contribute to finding solutions. The way we collect, analyse, interpret and share data can further reinforce racist mental models, practices and systems.

It is important to consider the historical context and how data, especially within science and medical research, has created and reinforced racist, sexist, ableist and other oppressive mental models which form the foundation of many policies and practices that continue to disadvantage some communities today. Race science focused on exploring the differences between races with the aim of creating an evidence base that suggested that the white race was superior to others. It reinforced the narrative of the evolutionary basis for health and social inequalities. This perspective is increasingly contested as there is greater recognition of the impact of racism, its links to the wider determinants and health outcomes. However, the legacy of these harmful mental models persists today. Early research peddled racist ideas such as Black people having higher thresholds for pain. This ideology has seeped into medical practice⁹ and can be seen in maternity care for Black women. In the UK pregnant women from Black and South Asian backgrounds are at greater risk of dying during pregnancy, childbirth and postnatally. They experience poorer levels of care^{10;11}. Black women are less likely to be offered pain relief during childbirth, which in turn leads to more traumatic deliveries and poorer maternal outcomes.

HOW TO REDUCE THE RISK OF CAUSING HARM

Making better use of data and the technological systems available to the NHS we could reduce gaps in patient data and address health inequalities. The health challenges we focus on, the way they are framed, the data we collect, and the actions that are taken in light of the evidence available, are often inherently biased. If we pay closer attention to systemic issues we will be better equipped to minimise their impact or prevent harm coming to Black and South Asian communities.

Collecting data

The data we choose to collect determines whose needs are visible. If a healthcare worker decides not to collect information from a patient and they are not represented in these datasets, it makes it difficult for services to understand their needs, the nature of the inequalities they face or to find solutions.

Mental health services in London wanted to know about young people from Black and Asian backgrounds access, experience and outcomes in services. The data available highlighted data quality on ethnicity was poor and it was not possible to draw any meaningful conclusions, nor was it clear about how services needed to improve. It is important for healthcare workers to encourage patients to share their health information and to complete forms as accurately as possible.

Data that is collected is often determined by senior leaders within healthcare, regulatory bodies and the government. This information contributes to regional and national databases which identify health trends within populations. While it is important for Black and South Asian people to be represented in these datasets, the standardised approach may miss important information that could help to provide a more holistic picture. It can make it difficult for local health systems to pick up emerging trends in their localities. For example, a local area may have large populations from a racially minoritised group whose ethnicity may not be routinely recorded (e.g. Black Latinx community). Therefore, they are more likely to be categorised as 'other', this can make it difficult to understand the community's specific needs when the data is analysed.

When collecting data for standardised datasets, local services should consider working with local communities to identify other data that could supplement these datasets and which could provide information relevant to the local and cultural context.

HOW TO REDUCE THE RISK OF CAUSING HARM

Within healthcare services and research some communities are hyper visible or over surveilled and others are invisible.



The menopausal experiences and outcomes of South Asian and Black women receive little attention, despite these communities often experiencing poorer health outcomes. Health services rarely collect data on women of colour who identify as lesbian, bisexual or queer and similarly there is an absence of information about their health experiences within the academic literature.

To ensure that health services are not introducing or replicating biases in their data they should collect data across the protected characteristics, consider the role of wider societal factors and make sure that patient information for racially minoritised groups is complete.

When collecting data from patients it is important to be aware who will have access to the information you are documenting. Within healthcare there has been a move towards sharing data between teams, departments and organisations to improve the quality of care people receive and to reduce the number of times they are required to tell their story. However, some

information could make people vulnerable and open to discrimination especially sensitive information such as mental health status, contact with the criminal justice system, immigration status etc. When recording information it is important to keep records factual and to ensure they are framed in a way that limits opportunities for misinterpretation or for the information to reinforce racist stereotypes.

Healthcare workers should be aware of the power imbalance between them and their patient or carer. Some people from Black and South Asian communities adhere to cultural norms that respect hierarchical structures within society. Therefore, challenging or criticising a health professional's advice or request may be considered inappropriate. People should not feel pressured in to giving or providing their data to be used and patients should be reminded that they have a choice.

Analysing data

When an analyst explores a dataset, they are often exploring the data to help answer questions that will assist services to make more informed decisions. However, the questions that are asked of the data may be biased and risk creating or reinforcing discriminatory narratives which then shape policy and practice. We may find that we focus on a particular area at the expense of

HOW TO REDUCE THE RISK OF CAUSING HARM

identifying other factors that are as or more important. When analysing data it is important to critique the assumptions that are made or explored and consider them within a historical, cultural, political and economic context.

Although we often consider the experiences of groups of people who share similar characteristics, we need to remember that diversity exists within these communities. As we exist within a society where power is inequitably distributed our social identities are often assigned privilege and disadvantage. These identities may intersect and shape our lived experiences of health and services. For example the experiences of a disabled Asian man may be different to an Asian man who does not have a disability. While they may share similar experiences on the basis of their shared race and gender, it is quite likely that a disabled person may have a very different experience. Similarly health services and health research often aggregate data on all racially minoritised groups into one category (e.g. BAME). A key challenge presented by this approach is that it masks the different experiences and outcomes between groups. This means that recommendations that may emerge may not be effective for all groups because of their differential experiences. Data on patient groups should be routinely analysed by race and where it is possible ethnicity should be considered at the intersection of the protected characteristics and class.

Sharing data

Once data has been analysed a decision needs to be made about how the insights are shared and used. If this is not considered carefully it may create and deepen inequalities. The practice of organisations withholding or concealing data to preserve the reputation of an organisation or service that may be failing Black or South Asian communities is a form of systemic racism. For example during the Covid-19 pandemic the government did not respond to Sheffield Hallam University's freedom of information request for data on all staff deaths as a result of Covid-19 from 1 March 2020 to 15 January 2021. The suppression of the data means that the unfair outcomes and discrimination remain misunderstood and are not addressed. At the same time the way in which data is shared can also reinforce stereotypes.

Therefore services need to be mindful about how the data they share might be received by people who may not have deep insight into the nature of inequalities or how one's exposure to institutional racism shapes their outcomes. For example, during the Covid-19 pandemic the way data was communicated about groups who were at increased risk led to an increase in discrimination experienced by racially minoritised groups. ¹³ When sharing data the narrative should contextualise the findings to highlight structural factors that contribute to the inequalities that have been observed.

HOW TO EARN TRUST AND BUILD CONFIDENCE

The harms that Black and South Asian communities have experienced at the hands of medicine, science and other statutory bodies have played a significant role in breaking trust and confidence in institutions (e.g. Tuskegee experiment in the USA, 14;15 death of Olaseni Lewis in hospital etc. 16;17). These incidents are framed as historical events that have little significance today, but many forget that many of these events occurred within people's living memories. In addition to this their lives continue to be limited by structural racism and other forms of oppression. While people from Black and South Asian communities may have shared experiences of oppression and distrust in institutions is common—there is diversity of thought and experience within and between groups and therefore feelings of distrust may not be universal.

Research suggests that institutions who are trusted create a sense of security which also promotes trust within communities. As an anchor institution the NHS needs to build bridges to gain the trust and confidence, in doing so this could provide benefits for the wider community they serve. To earn the trust of Black and South Asian communities NHS and its workforce must make a sustained commitment to transparency, accountability and to share power. To build the foundations for a trusting relationship we must all educate ourselves about the historical and current trauma racism causes people from Black and

South Asian communities' and support the NHS and partners to become antiracist.

To build trust the NHS and its workforce need to demonstrate to communities that:

- The NHS has a commitment to being anti-racist providing support as well as accountability mechanisms for staff who do not adhere to these principles
- The NHS is transparent, takes responsibility when patients are harmed in their care and puts measures in place to redress the harm caused
- Healthcare workers are aware of how data practices can cause harm and are equipped to prevent or minimise harm when collecting and using patient data
- NHS services have a clear understanding and are transparent about why data is collected and how it is used
- There is a commitment to involve them in shaping research, the design, and planning of healthcare services
- Their data is being used to improve services so it reflects local needs
- Systems are in place to keep their data secure and that there are consequences for individuals and organisations who compromise the security of patient health information
- Healthcare professionals are willing to take the time to support patients and carers to make informed decisions about their data

WHAT I CAN DO TO IMPROVE DATA PRACTICES



Everyone who works in healthcare can play a role in improving data quality. It is the small things that help to shift working culture . We have outlined a few things you can do depending upon your role to improve data practices:

As a frontline worker:

- Take the time to explain to patients why they are being asked to provide information and advise them on how it will be used and stored. Signpost them to other resources such as the Understanding Patient Data website
- Make an effort to collect equalities monitoring data and remember this supports efforts to reduce health inequalities
- Support patients to make an informed decision about whether they choose to share their data. Section 22-23 provides some guidance to help you have meaningful conversations with service users

- When colleagues present data or findings ask if they have analysed the data by ethnicity and other protected characteristics
- Develop relationships with Black and South Asian-led VCSE organisations who can help to facilitate conversations with the community about the role of patient data
- Provide timely access to interpretation and advocacy services to assist patients and carers to make informed decisions about how their data is collected and managed

WHAT I CAN DO TO IMPROVE DATA PRACTICES

As a decision maker:

- Train and support staff to understand the links between health inequality and systemic racism to enable them to have the confidence to engage with patients about how racism has shaped their experience within healthcare
- O2 Communicate clearly to the workforce the purpose of collecting data, how it will be used and stored
- Develop strategic partnerships with Black and South Asian-led voluntary, community and social enterprise organisations who can support with:
 - The engagement of communities in shaping local priorities for data collection
 - Developing a co-productive approach to data collection, analysis, interpretation and dissemination
 - Helping the health system to understand the nuances of racialised communities within health systems
 - The development of feedback mechanisms to keep Black and South Asian communities informed about how data is used to improve the health of their community
- Improve ethnicity data collection across all groups and pay particular attention to address missing data

- Ensure that all data is analysed by ethnicity and other protected characteristics and where possible explore the relationships between multiple protected characteristics and health outcomes
- Ensure that the findings and the actions that follow are shared with the staff who collect data so they are aware of the impact of their contribution
- Invest in the development of resources that are translated into community languages
- Demonstrate strong leadership and a genuine commitment to addressing health inequalities; implement accountability mechanisms to address poor performance and lack of engagement in inequalities work
- Incorporate information and guidance on holding conversations with racially minoritised groups about their data within compulsory data governance training, medical training and during staff inductions.
- When designing or introducing new software/technological solutions, coproduce then with staff, patients and carers
- Improve access to interpretation services, produce documentation in community languages and easy read formats
- Make better use of social media and other communication channels. For example, you may reach new and wider audiences by sharing information using videos and social media.

WHAT I CAN DO TO IMPROVE DATA PRACTICES





Birmingham and Solihull Mental Health Trust commissioned Catalyst 4 Change to undertake some research with Black communities to gain insight into their perspective on mental health and their experiences of services. They successfully recruited people to focus groups and interviews and passed the data to the trust.

The trust analysed the data and fedback their interpretation to the community. Their analysis suggested that Black people wanted better access to services, improved experiences and outcomes. However, communities felt the trust's interpretation did not reflect their perspective.

Black Thrive's Research Institute set up a data club and worked with community members to analyse the data. From a community perspective their narratives were expressing their desire to be be liberated and for services to take a more preventative role.

Increasingly voluntary organisations are being invited to collect data to support decision making within healthcare systems. This is a positive step to increasse the likelihood of engaging people who may not be engaged in contact with services. However, they should also be involved in analysig and interpreting the data. This will ensure that health systems build a well rounded picture of the potential opportunities and challenges.

WHAT I CAN DO TO IMPROVE DATA PRACTICES

As a policy maker:

- Work in partnership with communities and healthcare workers to place greater control of patient health information in the hands of the patient/service user
- Initiatives that aim to improve data collection should include processes to validate the accuracy of the data collected
- Consider how the learning from the <u>Patient and Carer Race Equality</u>
 <u>Framework</u> could support other services to improve data practices
 within physical health and social care services
- Explore options to expand ethnicity categories to enable service users to select an ethnicity that is aligned to how they self identify and a free text option for people to self define.



The Patient and Carer Race Equality Framework^{18;19;20} requires mental health trusts to explore their data to understand where inequalities exist for communities who are racialised as minorities.

As part of the programme they work with community partners, service users and carers to explore trends in the data. They work together and use the data to set priorities for service improvement. The data is shared openly and reviewed regularly to track progress and to hold the trust to account.

This process helps the trust to gain insight into the stories that sit behind the quantitative data they hold, so they have a better understanding of the communities needs. The transparency that is encouraged by the PCREF also helps to build the trust and confidence of communities involved in the process.

HOW TO SUPPORT PATIENTS TO UNDERSTAND HOW THEIR DATA IS COLLECTED AND USED

The infographic included provides information that will help to guide your conversations with patients from Black and South Asian communities about the benefits of sharing their data and how it is used. As practices across trusts and health organisations vary you should check that the organisation you work for is following best practice so you do not mislead patients. Visit NHS Digital's website²¹ for more detailed information about how patient data is managed within the NHS

We have included a few principles for you to keep in mind when holding conversations with patients and carers:



People have legitimate concerns about the negative consequences that may arise if they share their data



Encourage patients to ask questions. Be open when you don't know the answer to a question and make the effort to find the answer



Ensure patients have the option to say no without being made to feel that they are being difficult or unreasonable



Allow people to take their time to come to a decision. You may find that some people may refuse, but when given the opportunity to think about the pros and cons they change their mind

Black, South Asian and (In)visible:

Why patient data matters

Patient data for Black and South Asian communities is more likely to be missing or recorded incorrectly



What is patient data?

The NHS collects a lot of different types of data here are a few:



- Clinical data (Blood pressure, results from tests)
- Feedback on services (e.g. surveys, complaints etc.)
- Protected characteristics (e.g. age, ethnicity etc.)
- Social/cultural information (e.g employment status, religion

Why does it matter if data is collected from Black and South Asian communities?

Collecting patient data and using it well is a crucial step. It will ensure that people from these communities have equal opportunity to have good health, access to high quality and safe services and live full lives.



Better able to meet a patient's care needs



Helps to provide good quality and safe healthcare



Delivers services designed to meet the needs of the local population



Supports the development of new and improved approaches to treat and prevent illness or disease



Supports better policies which influence our access to healthcare



Helps services to understand how factors outside of the health service affect our health (e.g. debt, housing etc.)



Improves methods and efficiency in diagnosing illnesses





Concerns people may have about how their data is used



Sharing data with other agencies



Discrimination



Selling data to private companies

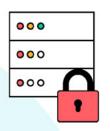


Data breaches

How is a patient's privacy protected?

It is a criminal offence to misuse patient data. NHS-specific guidelines are intended to ensure that confidential information is handled responsibly and in accordance with data protection laws. These guidelines specify that organisations must make every effort to:

Keep data secure. Use data that cannot identify individual patients whenever possible. Use data to benefit health and care. Not use data for marketing or insurance purposes (unless this is requested by the patient). Make it clear why and how data is being used.



Where is patient data stored?

A patient's data is collected during each contact with a health or social care service. This information may be stored as paper records or electronically.



Who has access to a patient's information? Once patient data has been anonymised or pseudonymised and provided criteria in 5 met



Non-profit research organisations



Commercial organisations



Government departments



Local Councils



Primary Care Networks



Integrated Care Systems

What if a patient does not want to share their information?



Some information is needed to provide their care. However, they can opt out from providing data that is used for research or to plan services by accessing NHS Digital's opt out service.



For more information



About patient data



NHS Digital Opt out service



Your right of access



Complain about data breaches

WHERE CAN I FIND MORE INFORMATION ABOUT HEALTH DATA?



Understanding Patient Data has produced a number of guides including:

- Why is it important to use patient data?
- What happens to Health Data: A guide to large datasets
- To learn more about how the NHS uses data visit NHS Digital

Ada Lovelace Institute focuses on research to ensure Artificial Intelligence works for everyone. https://www.adalovelaceinstitute.org/about/

The NHS Race & Health Observatory works to identify and tackle ethnic inequalities in health and care. https://www.nhsrho.org/'

The Algorithmic Justice League is an organisation established in the USA who use art and research to highlight the social implications and harms of artificial intelligence. https://www.ajl.org/about

Health Data Research UK is a national institute that aims to pull together the UK's health data to enable discoveries that improve people's lives. https://www.hdruk.ac.uk

GLOSSARY OF TERMS

Anchor institution: In a healthcare context anchor institutions are large organisations (e.g. NHS Trusts) that have a strong investment in their community and are likely to be a permanent feature in the local context. They have substantial resources and influence that could be used to to promote the health and welfare of their local community and address health inequities. For example, procurement services locally training, providing local employment, a space to support collective action etc.

Anonymised: Data is considered to be 'anonymised' if all identifying information (such as your name, address or NHS number) is completely removed. This means that your identity cannot be re-linked to the data in a straightforward way. Data that is treated in this way is no longer considered confidential and, generally speaking, does not fall within data protection laws. It is important to note that even when health data is anonymised, it may still be possible to find ways of identifying your individual, personal information. However, it would likely require special circumstances or effort. This effort may involve using other sources of information to narrow down the number of individuals that the data may be referring to. As a result, there is some debate about when data should be considered completely 'anonymous'.⁷

Antiracism: An active commitment to working against racial injustice and discrimination. To be anti-racist is to be an active part of the solution, rather than just a bystander. Part of the role as an anti-racist person is self-reflection and self-improvement.

Include racially minoritised in the glossary of terms: Racialised minorities: People from communities that have been minoritised through social processes of power and domination rather than just existing as statistical minorities within a population. These groups may be described as ethnic minority groups, Black, Asian and Minority Ethnic (BAME) as well as other terms.

Intersectionality: The concept of intersectionality describes the ways in which systems of inequality based on ones social identities (e.g. gender, race, ethnicity, sexual orientation, gender identity, disability, class) and other forms of discrimination "intersect" to shape ones experience of privilege and or marginalisation.

People of Colour: A term commonly used in the USA and Canada to describe individuals and groups of people who are non-white.

Protected characteristics: Social identities that are protected against discrimination under the Equality Act 2010. These include age, disability, gender reassignment. marriage and civil partnership, pregnancy and maternity, race, religion or belief, sex and sexual orientation

Pseudonymised: Identifying information within a data record (name, address, etc.) has been replaced with a distinctive marker.

Racism: 'Racism is an organized societal system in which the dominant White group categorizes individuals into "races" and devalues, disempowers, and differentially allocates resources to ethnic groups considered to be inferior1. The ideology of inferiority permeates societal systems and institutions ensuring that racism is not limited to individual beliefs and behaviors1. Racism operates through institutional, interpersonal and cultural pathways'9

VCSE: An acronym used to represent the Voluntary, Community and Social Enterprise sector.

REFERENCES

- Understanding Patient Data. [Internet]. What happens to health data [cited 2002 Nov 28]. Available from: https://understandingpatientdata.org.uk/what-happens-health-data-guide-large-datasets
- Creary, N., Amevenu, D., Shivji, J. and Henry C. (2022), What it means to be seen: Closing gaps in patient data for Black and South Asian Communities. 2022 [cited 2022 Dec 13] London: Liberating Knowledge. Available from: https://understandingpatientdata.org.uk/ insights-black-south-asian-people-patient-data
- Scobie S, Spencer, J, Raleigh V. Ethnicity coding in English health service datasets. London: Nuffield Trust; [Internet]. 2021 [cited 2022 Nov 28]. Available from: https://www.nuffieldtrust.org.uk/files/2021-06/1622731816_nuffield-trust-ethnicity-coding-web.pdf
- 4. Naumova EN. Public health inequalities, structural missingness, and digital revolution: time to question assumptions. J Public Health Policy [Internet] 2021 [cited 2022 Nov 28]; 42(4):531-535. Available from doi:10.1057/s41271-021-00312-y
- Adamson AS, Smith A. Machine Learning and Health Care Disparities in Dermatology. JAMA Dermatol [Internet]. 2018 [cited 2022 Dec 08]; 154(11):1247-1248. Available from: doi: 10.1001/jamadermatol.2018.2348. PMID: 30073260.
- Knight HE, Deeny SR, Dreyer K, Engmann J, Mackintosh M, Raza S, Stafford M, Tesfaye R, Steventon A. Challenging racism in the use of health data. Lancet Digit Health [Internet]. 2021 Mar;3(3):144-146. Available from: doi: 10.1016/S2589-7500(21)00019-4.
- 7. Williams, D. R. and Etkins, O. S. (2021) Racism and Mental Health, World Psychiatry [Internet] 2021 [cited 2022 Dec 08]; 20 (2),pp. 194-195. Available at: https://doi.org/10.1002/wps.20845

- 8. Equality and Human Rights Commission. Your rights under the Equality Act 2010 [Internet]. [cited 2022 Nov 28]. Available from: https://www.equalityhumanrights.com/en/advice-and-guidance/your-rights-under-equality-act-2010#:~:text=Under%20 the%20Equality%20Act%2C%20there%20are%20nine%20protected,differences%20 depending%20on%20which%20protected%20characteristic%20you%20have.
- Hoffman KM, Trawalter S, Axt JR, Oliver MN. Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. Proc Natl Acad Sci U S A. [Internet] 2016 [cited 2022 Nov 28]; 113(16):4296-301. Available from: doi: 10.1073/pnas.1516047113.
- Knight, M., Bunch, K., Tuffnell, D., Shakespeare, J., Kotnis, R., Kenyon, S., et al. (Eds.).
 MBRRACE-UK: Saving lives, improving Mothers' Care 2020: Lessons to inform maternity care from the UK and Ireland confidential enquiries in maternal death and morbidity 2016–18. United Kingdom; National Perinatal Epidemiology Unit, University of Oxford; 2020
- 11. MacLellan, J., Collins, S., Myatt, M., Pope, C., Knighton, W., and Rai, T. (2022) Black, Asian and minority ethnic women's experiences of maternity services in the UK: A qualitative evidence synthesis. Journal of Advanced Nursing [Internet] 2022 [cited 2022 Nov 28]; 78(7), pp. 2175-2190. Available from: doi: 10.1111/jan.15233
- 12. Ramamurthy A. Government accused of withholding ethnicity data relating to Covid-19 deaths amongst NHS workers. Available from: https://www.shu.ac.uk/news/all-articles/latest-news/nursing-narratives-foi-ethnicity-data
- 13. The United Nations Human Rights Office of the High Commissioner. Racial discrimination in the context of the Covid-19 crisis. [Internet] 2020 [cited 2022 Dec 08] Available at: https://www.ohchr.org/sites/default/files/Documents/Issues/Racism/COVID-19_and_Racial_Discrimination.pdf

REFERENCES

- 14. Centers for Disease Control and Prevention. The Syphilis Study at Tuskegee Timeline [Internet]2022 [updated 2022 Dec 05; cited 2022 Dec 08]. Available from: https://www.cdc.gov/tuskegee/timeline.htm
- 15. Vonderlehr, R.A., Clark, T., Wenger, O.C., Heller, J.R., Untreated Syphilis in the Male Negro, Journal of Venereal Disease Information. 1936, 17:260-265.
- 16. GOV.UK. New law to prevent use of force in mental health settings [Internet] 2021 [2021 Dec 07, cited 2022 Nov 28]. Available from: https://www.gov.uk/government/news/new-law-to-prevent-use-of-force-in-mental-health-settings
- 17. Justice for Seni. The Olaseni Lewis Campaign for Justice and Change [Internet] 2020 [cited 2022 Nov 28]. Available from: https://www.justiceforseni.com/about/
- 18. Black Thrive Global [Internet]. What we do; 2021 [cited 2002 Nov 28]. Available from: https://blackthrive.org/what-we-do/
- Black Thrive Global. Black Thrive Impact Report 2021-22. [Internet]. 2022 [cited 2022 Nov 28]. Available from: https://blackthrive.org/wp-content/uploads/2022/10/Black-Thrive-Impact-Report_2021-2022_FINAL.pdf
- 20. NHS England. Advancing mental health equalities. [Internet]. [cited 2022 December 05] Available from: https://www.england.nhs.uk/mental-health/advancing-mental-health-equalities/
- 21. NHS Digital. Data. [cited 2022 December 08]. Available from: https://digital.nhs.uk/data

