



# Patient Data

## Finding the best set of words to use

Summary of findings

March 2017

# Contents

- Introduction and explanation of the process
- Detailed findings for each term
- Summary of all proposed terms
- Appendix: all words tested, exercises

# Introduction

The **current language landscape** around the use of patient data in care, treatment and research is difficult, complex and confusing. And current attempts to come up with alternatives have fallen short.

This acts as a **significant barrier to having open discussions** with the public about the use of data in ways that **can build both understanding and trust**.

Understanding Patient Data commissioned Good Business to conduct a creative development and research process to come up with and test **a set of words** (together with visual/ graphic representation) **which are simple, clear and accurate** to help build trust and understanding.



## Our project question and focus

**What is the best vocabulary to use to talk about the use of data for care, treatment and research?**

### *Key areas to cover*

**Overarching term for this space**

**Uses of data**

**Nature of data (level of identifiability)**

# What we did to answer this question

## 1. CREATIVE WORKSHOP

Ran a **creative language workshop** to generate possible alternative terms

## 2. EXPERT REVIEW

Explored alternative terms with **stakeholders and experts**

## 3. FOCUS GROUPS

Ran **12 focus groups with the public and health professionals** to explore terms further

Process designed to develop, refine and test a set of words that work for everyone – professional and public

- Keen to ensure we don't reinvent the wheel, add confusion to the landscape or try to fix terms that aren't broken.
- Recognise it's unlikely to be feasible to come up with a set of terms that all stakeholders strongly endorse: looking for a pragmatic solution that helps make progress and aids clarity.

# The first step was a creative language workshop

We gathered a series of language experts from different spheres to collectively explore the current vocabulary and come up with possible alternate words to use for the terms. We also challenged them to come up with ideas for visual imagery which would help people understand the terms.

## Participants

Range of experts from different areas (outside of health):

- Verbal brand consultant
- Data journalist
- Science writer
- Speechwriter
- Technical writer
- Copywriter
- Linguist
- Writer

## Process

Series of structured group exercises and ideation sessions designed to fuel the creative process while also sense checking ideas. Participants considered current terms too – and were instructed ‘if it’s not broken don’t try and fix it!’

# The second step was an expert review of the words that came out of the creative workshop

We **conducted phone interviews with experts** who are close to discussions about patient data in some way to **explore their views on the alternative terms** generated at the workshop. We wanted to **canvass their views**, and **understand any no-go areas** before testing words with focus groups.

## Participants

Range of experts from different areas:

- NHS
- Department of Health
- National Data Guardian office
- Information Governance Alliance
- Connectedhealthcities
- MedConfidential
- ALSPAC (Avon Longitudinal Study of Parents and Children)

## Process

We sent the experts the working framework in advance of the call and then went through each term to explore pros and cons – as well as any new alternatives

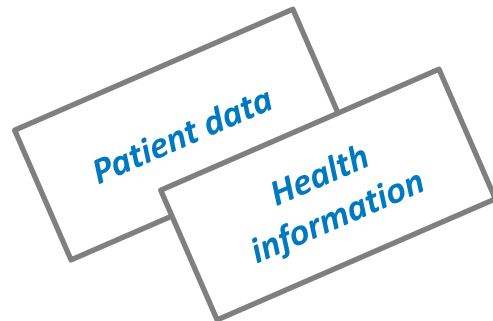
# The final stage of the process was a series of focus groups

12 x 60 minute groups with up to 12 respondents in each group

Number	Group	Date	Location
1	Nurses	1/11	Central London
2	Practice Managers		
3	GPs		
4	Public, nationally representative	2/11	Central London
5			
6			
7	Public, nationally representative	2/11	Suburban
8			
9			
10	Practice managers	3/11	Suburban
11	Nurses		
12	Hospital specialists/dentists		



# In the groups we used group discussions, exercises and private capture to test reactions and explore challenges



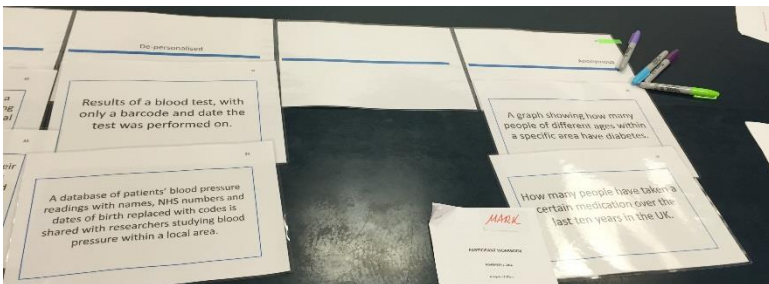
**TERM 2 Definition:** The use of a person's health information for their own diagnosis, care and treatment by health and social care professionals.

- Individual care
- Personal care
- Direct care
- \_\_\_\_\_

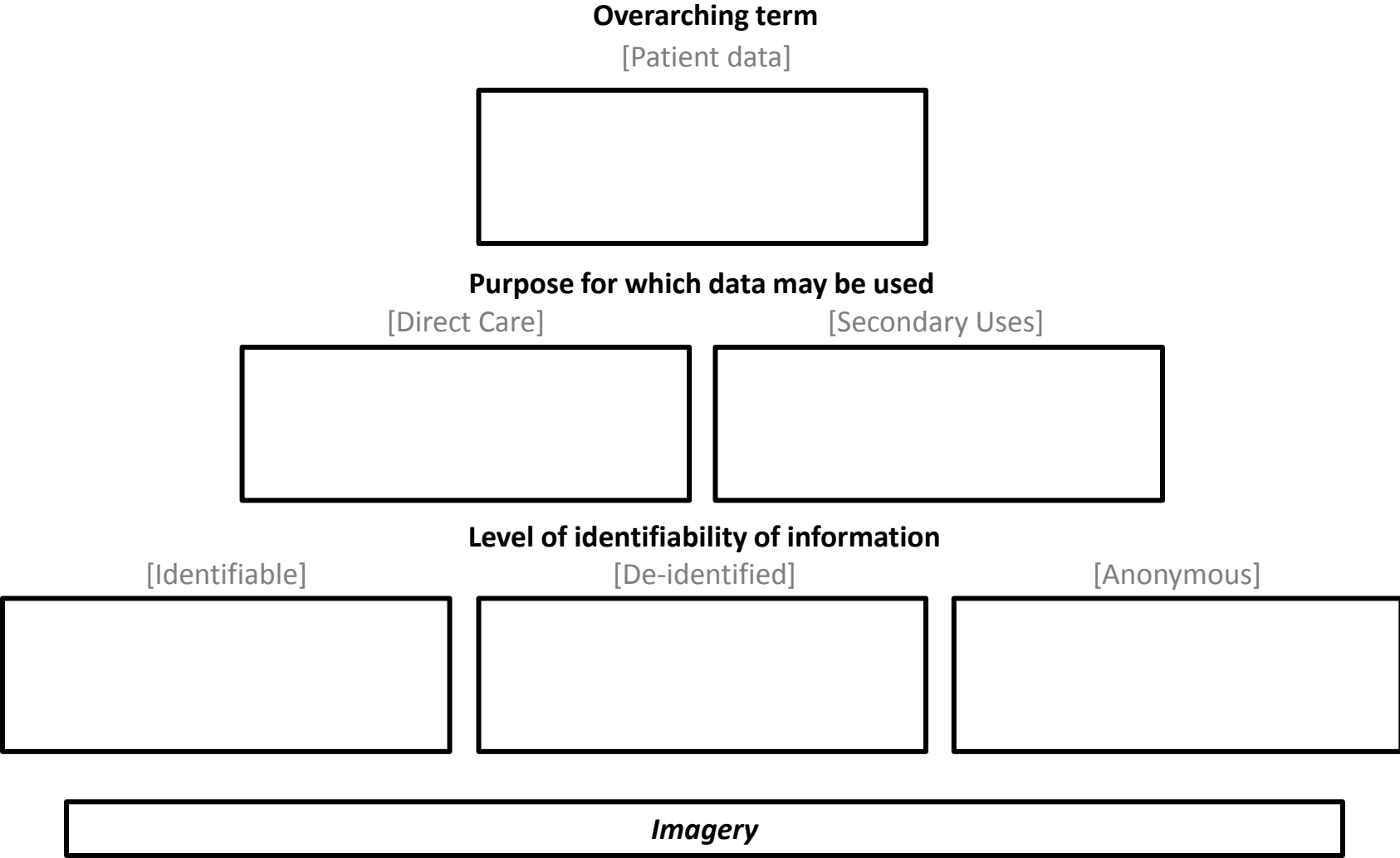
Spontaneous reactions to terms – what do people think they mean? Does this match the definition?

Participant workbooks – self completion to select preferred term (or offer alternatives)

Exercises to test understanding and explore points of tension



# Throughout the process we used this framework to structure the areas of exploration



## Detailed findings for each term

# Health information came out of the creative workshop, though subsequent discussions with experts raised questions

## Creative workshop findings

Preferred term:

**Health information**

Other suggested terms:

*Health details; Patient health information; Health information about you*

- Many felt that as the term needs to cover both numbers and codes and written information (e.g. doctor's notes) **information is better than data**
- Emphasis on the fact that the term needs to make it very clear which area this is in – so including a word like **medical, patient, or health** is key, and of these health seen to be the most personal and approachable – this led to *health information*
- Balancing **accuracy, understanding and feeling** can be tricky – terms like *patient records* and *patient history* are easily understood but people might feel more 'precious' about them, also suggests their records would be used in their entirety

*Information* includes everything, unlike *data* which just brings to mind numbers and codes

The word *record* makes me think they're taking my details for some other purpose

## Input from expert discussions

*Health information* could suggest information for you rather than about you (e.g. an information leaflet about a condition). Might not naturally include social care information.

# We took several alternatives to the focus groups – and found *patient data* and *patient health info* rose to the top

## Focus group findings

Preferred terms:

***Patient data***

***Patient health info***

Terms explored:

***Health information/details/data; Citizen health information; Health information about you; Your health information; Patient health information; Information from patient records***

- *Patient data* was **the favoured term by a narrow margin** – seen as the easiest term to ‘get’ and a good match to the definition, though also a bit cold and impersonal and doesn’t make people think of softer forms of data (e.g. notes)
- ***Patient health information*** was also liked by many, who feel it ‘says it like it is’, and includes all forms of data
- People **don’t naturally think either term covers social care information**, but when questioned around this **they accept it** - the fit is slightly better with ‘patient health information’ as it’s broader
- When it came to *health information* most went to information **for them (e.g. a fact sheet)**
- Discussions around other terms revealed a need to strike a balance between being **easy to understand but not patronising**, and **technical but not authoritative**
- The use of ‘your’ or ‘about you’ had mixed reactions – some thought it made the terms clearer, others thought it was **too individualistic** and pointing the finger at them

[*Patient health information*] is specific, and it’s got to be made idiot-proof for when we’re dealing with patients. (HCP)

I originally thought this [*Health information*] was information about how I can be healthy. (Public)

[*Your health information*] sounds like someone’s talking down to me. (Public)

# Proposed terms

## Overarching term

[Patient data]

**Patient data**  
**(Patient health information)**

*It may be that there is no single answer for this category - patient data is the simplest term to use but has drawbacks, and 'patient health information' houses a broad range of info more comfortably, so a dual approach may make sense*

## Purpose for which data may be used

[Direct Care]

[Secondary Uses]

## Level of identifiability of information

[Identifiable]

[De-identified]

[Anonymous]

**Imagery**

# The creative workshop came up with alternatives to both the current terms, which experts broadly supported

## Creative workshop findings

Direct care		Secondary uses	
Preferred term: <i>Individual care</i>	Alternative term: <i>Personal care</i>	Preferred term: <i>Improving health, care and services</i>	Alternative term: <i>Research, planning and development</i>

- Experts felt *direct care* sounded a bit **'ominous'** and didn't make what would be covered obvious, though *care* was seen as a good word to use, it is **warm**, and covers more than *treatment* – hence 'individual care'
- **Secondary uses was seen as removed and uninformative** – hard to understand what it means. Experts felt being clearer about the benefits of the use would help people 'get it' which led to 'improving health, care and services'
- Discussions also considered terms such as *societal* or *universal* as they communicated that the information is used for the greater good
- Some felt it would be useful to explain **how the data would be used** to help reassure

'Individual', when used in context, speaks volumes.  
(Language expert)

People often don't feel that 'research' relates to them - we need to show how these secondary uses relate to the individual (Language expert)

### Input from expert discussions

High levels of positivity around *individual care* as a real improvement on *direct care* and favourable reaction to *improving health, care and services*

# Individual care worked very well in the groups, but a modification to *improving health, care and services* came out

## Focus group findings

### Direct care

Preferred term:  
***Individual care***

Other term explored:  
***Personal care***

### Secondary uses

Preferred term:  
***Improving health, care and services through research and planning***

Other terms explored:  
***Improving health, care and services for everyone; Research, planning and development; Universal care***

- *Individual care* was the preferred option by almost all participants and was spontaneously interpreted as intended. *Personal care* was interpreted by many to be about **sanitary care/hygiene**
- *Improving health, care and services* was also interpreted as intended, but sounded **a bit like politician-speak** to some, and for others raised questions around how the data would actually be used – **so *improving health, care and services through research and planning* was preferred**
- *Improving health, care and services for everyone* was seen as a strength for some, who like the **feeling of altruism** that this gave to the term – others were more cynical about the system and thought this was misleading
- *Research, planning and development* was seen to be **dry and ‘cold’ – and not health specific**

It makes me think of the ladies who come round to look after elderly people. (Public)

[*Improving health, care and services for everyone*] “for everyone” – this is not true, only in a utopia. We don't live in an ideal world. “Changing” rather than “improving” .  
(HCP)

Sounds like house planning, doesn't sound like it relates to patients. (Public)



## We used an exercise to explore whether people would expect specific examples of use to fit under ‘improving health, care and services through research and planning’

Most examples were seen to fit – though the financial aspect raised questions for some

Number of groups who placed example in each category			
	Yes	No	Unsure
A study of over-50 year olds as they age, conducted to determine whether lifestyle is linked to Alzheimer’s disease.	12	0	0
A GP surgery analysing patient records to determine how many flu vaccinations they should order for the next winter.	12	0	0
An NHS-commissioned study of how many antibiotics are prescribed unnecessarily by GPs.	11	1	0
A private analytics company working in partnership with the NHS to look at how best to provide kidney dialysis services to get the best outcomes.	10	0	2
A pharmaceutical company uses patients’ health information - with identifiable details removed - to determine if there are any long-term effects from a drug that it makes.	9	1	2
A group of GP practices review health info from their practices relating to arthritis prescriptions. GPs agree to only prescribe 4 types of drugs, reducing their drugs bill by £600,000.	7	1	4

For full wording of each example, see appendix.

Makes sense because our data is gathered to help with distribution. (Public)

Pharma and care don’t go together in my head. Pharma means money. (Public)

It’s misleading, a patient would assume any savings would go directly back into healthcare. (HCP)

# Proposed terms

Overarching term

[Patient data]



Purpose for which data may be used

[Direct Care]

[Secondary Uses]



Level of identifiability of information

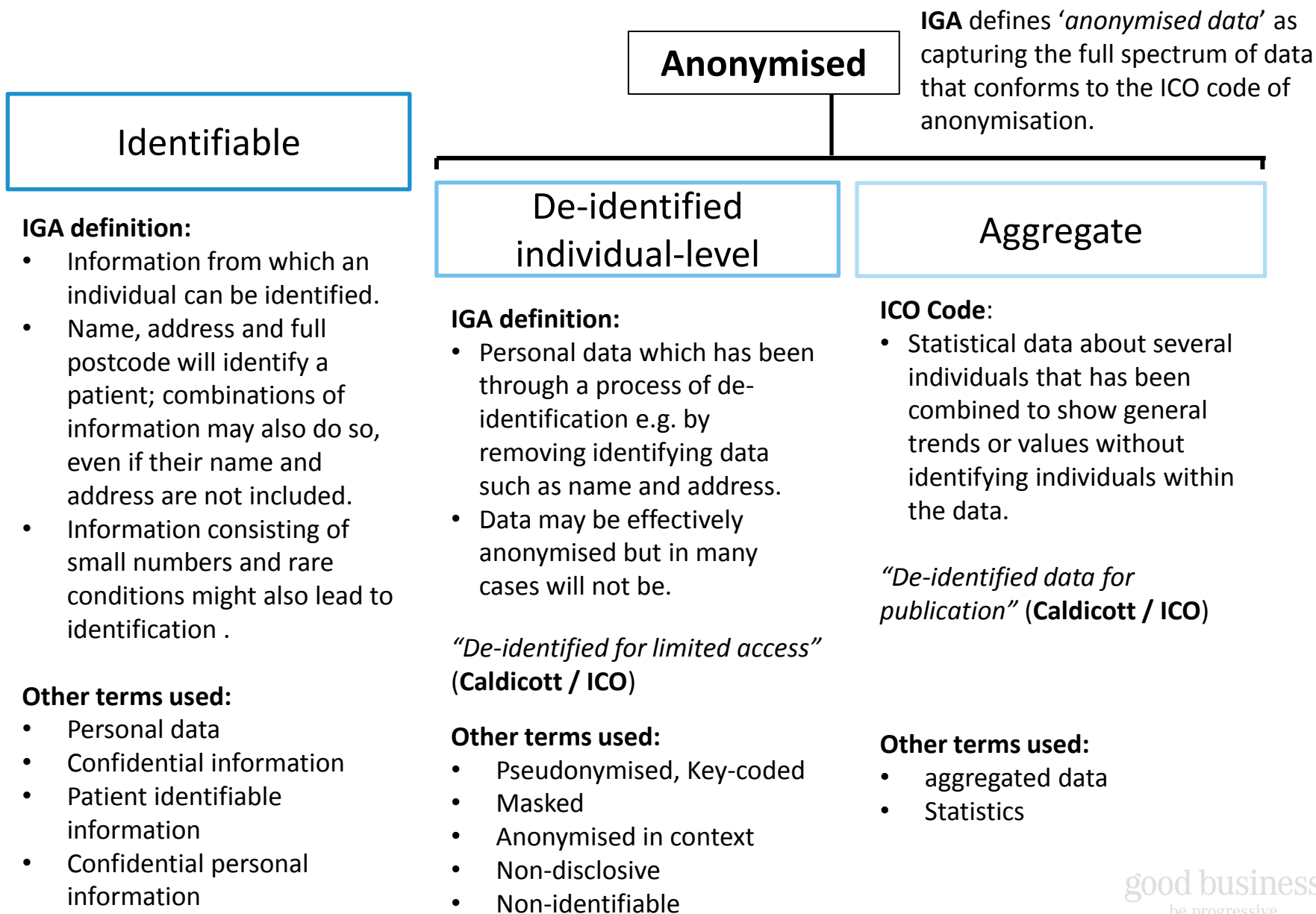
[Identifiable]

[De-identified]

[Anonymous]



# Levels of identifiability



## This area was (unsurprisingly) the least straightforward – but the creative workshop generated a couple of new terms

### Creative workshop findings

Identifiable	De-identified	Anonymous
Preferred term: <b><i>Personally-identifiable</i></b>	Preferred term: <b><i>De-personalised</i></b>	Preferred term: <b><i>Anonymous</i></b>
Other suggested terms: <b><i>Personal; Identifiable</i></b>	Other suggested terms: <b><i>De-identified</i></b>	Other suggested terms: <b><i>Anonymous grouped; Generalised</i></b>

- Workshop discussions considered taking a ‘**system approach**’ which forces another layer of explanation (e.g. A, B, C) as it’s too complex to cover in one word
- Feeling that making the link with the individual was key - which led to *personally-identifiable*
- Needs to be clear that de-identified data could be linked back – using the ‘*de*’ prefix in *de-personalised* helps with this – though without bringing up too many questions
- There was general consensus that *anonymous* works quite well as is, so no need to come up with something new

The problem with using numbers to describe the identifiability is that people don’t necessarily know the hierarchy (Language expert)

### Input from expert discussions

*High levels of positivity around de-personalised, and personally-identifiable also well received. Some concern that anonymous doesn’t get at the grouped nature of the data in this context.*

## *Personally-identifiable* and *de-personalised* worked well in the groups, *anonymous* liked but didn't encompass the group element

### Focus group findings

Identifiable	De-identified	Anonymous
Preferred term: <b><i>Personally-identifiable</i></b> Other terms explored: <b><i>Personal; Identifiable</i></b>	Preferred term: <b><i>De-personalised</i></b> Other terms explored: <b><i>De-identified</i></b>	Preferred term: <b><i>Anonymous</i></b> Other terms explored: <b><i>Anonymous grouped; Generalised</i></b>

- *Personally-identifiable* emerged as the clear favourite – seen as a good fit with the definition and spontaneously interpreted as intended. *Personal information* didn't always make people think it included information on their health (could just be name and address)
- *De-personalised* was favoured by most and easily understood, although a few felt it sounded **negative** or **'not human'**
- *Anonymous* works at a top level for most people – most spontaneously think it means data that can never be linked to an individual. Though they do not spontaneously link it to grouped data, rather that all details have been completely taken away / were never collected

A more understandable term [de-personalised] because it doesn't mislead someone to think it could NEVER be linked back to you. (Public)

It doesn't bring together the "groupedness" of it. If it was an amalgamation of data it would be clearer. (HCP)

## We used a mapping exercise to explore understanding – participants were asked to place examples along a ‘spectrum of identifiability’

The exercise revealed relatively high levels of understanding, though bar codes and databases were confusing

Number of groups who placed example in each category

	Personally identifiable	De-personalised			Anonymous	Unsure
The number of people prescribed a certain medication over 10 years.	0	0	0	0	12	0
Graph of diabetes rates within a local area.	0	0	0	1	11	0
A patient’s vaccination history from GP with NHS number.	11	1	0	0	0	0
GP reporting side-effects of a drug, including age and gender.	0	4	7	1	0	0
NHS database about A&E admissions with identifying details removed.	0	1	5	3	2	1
Results of a blood test with only a barcode and date of test attached	0	1	5	3	1	2
A database of patients’ blood pressure readings with names, NHS numbers and dates of birth replaced with codes	1	2	5	1	3	0

*Personally identifiable*

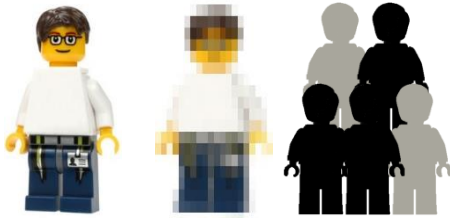
*De-personalised*

*Anonymous*



The different boxes under ‘de-personalised’ in the table reflect where participants placed the example on the spectrum – all fit under the de-personalised category

## We also explored some imagery - the concept of the picture/ pixelated picture/ silhouette was very well received



### Level of identifiability

- Instantly comprehensible to the majority and most felt it would be helpful - either in 'getting it' (the public) or explaining it (HCPs)
- Many felt it was best on its own, without the security element



### Security

- Most understood this to relate to security, though because security is linked to identifiability too, some struggled with the extra dimension (of the data environment)
- Generally, the shield/padlock combination was confusing and most preferred **padlocks alone**



### Composite

- To some people it was helpful to have the two concepts combined – and a few HCPs said it would be useful if they were talking this through with patients
- But for others it was too confusing and required too much explanation

# Proposed terms

## Overarching term

[Patient data]

**Patient data**

## Purpose for which data may be used

[Direct Care]

[Secondary Uses]

**Individual care**

**Improving health, care and services through research and planning**

## Level of identifiability of information

[Identifiable]

[De-identified]

[Anonymous]

**Personally-identifiable**

**De-personalised**

**Anonymous grouped**

**Imagery: clear image, pixelated image, grouped silhouette**



## Summary of all proposed terms

# A final set of terms to consider – which combines input from all stages of the process

**Overarching term**

[Patient data]



**Purpose for which data may be used**

[Direct Care]

[Secondary Uses]



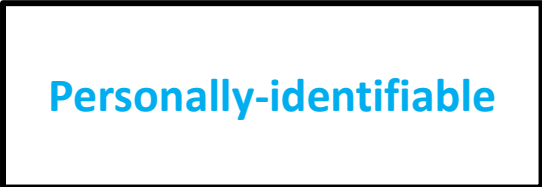
**Level of identifiability of information**

**Anonymised**

[Identifiable]

[De-identified]

[Anonymous]



**Imagery:** clear image, pixelated image, grouped silhouette

# Appendix

Good Business Ltd  
25 Gerrard St  
London, W1D 6JL, UK

Registered in England  
Company no. 3561306  
VAT number 701362381

For further information contact:  
Claire Jost  
Managing Director  
+44 (0) 207 494 0565  
[claire@goodbusiness.co.uk](mailto:claire@goodbusiness.co.uk)

## Appendix: Definitions for each term

Original Term	Definition used for testing new terms
Patient data	Information and data extracted from patient records that can be used for a wide variety of purposes. Examples of this information include details of medical conditions; notes recorded by healthcare professionals; and personal details such as NHS number and date of birth. This term could refer to the whole record, or just a part of it.
Direct care	The use of a person's health information for their own diagnosis, care and treatment by health and social care professionals.
Purposes beyond direct care	The use of a person's health information for purposes beyond their own diagnosis, care and treatment e.g. for medical research, public health research and monitoring, health service planning, and education and training.
Identifiable	Details from patient records that can be linked to a specific person because they include an NHS number, date of birth, postcode or any other piece of information that identifies the person. This information is stored in a highly secure way.
De-identified	Health information that cannot easily be linked to a specific person because the NHS number, date of birth, postcode and/or any other piece of information that identifies the individual has been removed, disguised or encrypted. Although this information cannot easily be linked back to you, with enough time and the right resources, the person could be identified.
Anonymous	Health information from many people that has been combined together to show general trends and therefore could not be linked to a specific person. As it only relates to large groups of people, it can't be linked back to a single person, and so it has fewer security measures attached to it.

## Tick tallies from the focus groups

# Overarching term

Term	Ticks	Crosses
<i>Patient data</i>	49	20
<i>Patient health information</i>	40	22
<i>Health information about you*</i>	21	55
<i>Your health information*</i>	14	33
<i>Information from medical records*</i>	12	17
<i>Health information</i>	7	56
<i>Health data</i>	6	47
<i>Health details</i>	5	46
<i>Citizen health information*</i>	2	47

***Patient data*** also the preferred term from the patient panel

\* Terms not shown in all groups

## Expert input

Input mainly around *health information* which was the favoured term from the workshop – questions were raised around whether it a) could encompass social care and b) made clear it was information about a person or individual rather than for them (e.g. information leaflet about a condition)

# Uses of data

Term	Ticks	Crosses
<i>Individual care</i>	95	7
<i>Personal care</i>	22	53
<i>Direct care</i>	4	74
<hr/>		
<i>Improving health, care and services through research and planning</i>	59	15
<i>Improving health, care and services for everyone</i>	36	21
<i>Improving health, care and services</i>	33	26
<i>Research, planning and development</i>	16	40
<i>Purposes beyond direct care</i>	2	75
<i>Universal care*</i>	1	43

***Individual care*** also the preferred term from the patient panel for the first category, with an equal split between the three '*improving...*' options for the second

\* Terms not shown in all groups

## Expert input

*Individual care* was very well received and was seen as a real improvement on *direct care*. There was a question whether it mattered that direct care had a statutory definition.

Input on second category mainly around *improving health, care and services* as the favoured term for the workshop – generally a very positive reaction to this.

# Level of identifiability

Term	Ticks	Crosses
<i>Personally-identifiable information</i>	67	26
<i>Personal information</i>	48	23
<i>Identifiable information</i>	13	33
<hr/>		
<i>De-personalised information</i>	97	10
<i>De-identified information</i>	13	45
<hr/>		
<i>Anonymous information</i>	56	14
<i>Anonymous grouped information</i>	40	26
<i>Anonymous pooled information</i>	21	31
<i>Generalised information</i>	12	38
<i>Generalised anonymous information</i>	10	33

## Expert input

Enthusiasm around *personally-identifiable* and *de-personalised* from most, though one voiced a reservation that *de-personalised* sounds less human. Also a concern from one around the breadth of the *de-personalised* category, and acknowledgement that there is a mental health condition known as depersonalisation-derealisation Disorder.

*Anonymous* raised a few more concerns - some felt it didn't get at the grouped nature of this data in this context, one also felt this data could never be truly anonymous in all cases.

*Personally identifiable* and *de-personalised* also the preferred terms from the patient panel for the first two categories, with a small majority for *anonymous pooled* for the third.



## **Exercise 1**

**Exploring what fits under *improving health, care and services through research and planning***

# Examples of secondary care

Original Term	Examples
Purposes beyond direct care	A pharmaceutical company uses patients' health information - with identifiable details removed - to determine if there are any long-term effects from a drug that it makes.
	A private analytics company working in partnership with the NHS to look at how best to provide kidney dialysis services to get the best outcomes.
	A GP surgery analysing patient records to determine how many flu vaccinations they should order for the next winter.
	An NHS-commissioned study of how many antibiotics are prescribed unnecessarily by GPs.
	A study of over-50 year olds as they age, conducted to determine whether lifestyle is linked to Alzheimer's disease.
	A group of GP practices prescribes medication for patients with arthritis. A review of health information shows that many types of anti-inflammatory drugs are being prescribed, including expensive ones that have the same result for the patient as cheaper options. The GP practices agree to only prescribe four types of the drugs, reducing their drugs bill by £600,000 for their area.

## Findings – most examples were seen to fit under the term

Participants discussed the example provided and reached a group decision as to whether it fit under *improving health, care and services through research and planning* or not

Example provided	Number of groups		
	Yes	No	Unsure
A study of over-50 year olds as they age, conducted to determine whether lifestyle is linked to Alzheimer's disease.	12	0	0
A GP surgery analysing patient records to determine how many flu vaccinations they should order for the next winter.	12	0	0
An NHS-commissioned study of how many antibiotics are prescribed unnecessarily by GPs.	11	1	0
A private analytics company working in partnership with the NHS to look at how best to provide kidney dialysis services to get the best outcomes.	10	0	2
A pharmaceutical company uses patients' health information - with identifiable details removed - to determine if there are any long-term effects from a drug that it makes.	9	1	2
A group of GP practices review health info from their practices relating to arthritis prescriptions. GPs agree to only prescribe 4 types of drugs, reducing their drugs bill by £600,000.	7	1	4

## **Exercise 2**

**Exploring levels of understanding around identifiability by mapping examples**



# Most examples were placed in the correct position, though databases and barcodes confused people

Example provided	Number of groups					
	<i>Personally identifiable</i>		<i>De-personalised</i>		<i>Anonymous</i>	Unsure
The number of people prescribed a certain medication over 10 years.	0	0	0	0	12	0
Graph of diabetes rates within a local area.	0	0	0	1	11	0
A patient's vaccination history from GP with NHS number.	11	1	0	0	0	0
GP reporting side-effects of a drug, including age and gender.	0	4	7	1	0	0
NHS database about A&E admissions with identifying details removed.	0	1	5	3	2	1
Results of a blood test with only a barcode and date of test attached	0	1	5	3	1	2
A database of patients' blood pressure readings with names, NHS numbers and dates of birth replaced with codes is shared with researchers studying blood pressure within a local area.	1	2	5	1	3	0

# Suggested terms from the creative workshop

## Overarching term

*[patient data]*

**Health information**

*Health details*

Terms in grey: original terms  
Terms in blue bold: preferred choice  
Terms in blue: alternative options

## Purpose for which data may be used

*[direct care]*

**Individual care**

*Personal care*

*[secondary uses]*

**Improving health, care and services**

*Research, planning and development*

*Universal care*

## Level of identifiability of information

*[identifiable]*

**Personally identifiable**

*Personal*

*[de-identified]*

**De-personalised**

*[anonymous]*

**Anonymous**

*Generalised*