THE UPPER TRIBUNAL
(ADMINISTRATIVE APPEALS CHAMBER)

UPPER TRIBUNAL CASE NO: GIA/0136/2021
[2021] UKUT 192 (AAC)
NHS BUSINESS SERVICES AUTHORITY V INFORMATION COMMISSIONER AND SPIVACK

Decided following an oral hearing on 6 July 2021

Representatives

Appellant	Robin Hopkins of counsel
Information Commissioner	Eric Metcalfe of counsel
Mr Spivack	Took no part

DECISION OF UPPER TRIBUNAL JUDGE JACOBS

On appeal from the First-tier Tribunal (General Regulatory Chamber)

Reference: EA/2019/0407
Decision date: 9 November 2020

The decision of the First-tier Tribunal did not involve the making of an error on a point of law under section 12 of the Tribunals, Courts and Enforcement Act 2007.

REASONS FOR DECISION

A. The request for information and its fate

1. Mr Spivack is interested in prescriptions for Stiripentol (also known as Diacomit), which is a drug prescribed to children for epilepsy. He made regular requests, beginning in 2013, to NHSBSA under the Freedom of Information Act 2000 asking for information. This case concerns the request made in January 2019, relating in effect to March to November 2018 inclusive. Specifically, he wanted 'A list of dispensaries (by dispensary code only), which have dispensed the following products, along with prescription items, cost and quantity for each dispenser by individual month'. This information had always been provided, but this time NHSBSA

refused to provide the name and location of the dispenser for all but two of the entries in its records. It explained that it had withheld information where a dispenser had supplied fewer than five items 'because patients could be identified, when combined with other information that may be in the public domain or reasonably available.' All the other information requested was supplied for all the entries.

2. On complaint under section 50 of the 2000 Act, the Information Commissioner decided that NHSBSA should have provided the withheld information and required it to do so. The First-tier Tribunal dismissed NHSBSA's appeal, but gave permission to appeal to the Upper Tribunal.

B. What I have to decide

3. I have to decide whether any person was identifiable from the data withheld by NHSBSA when taken together with other information by someone who was motivated to identify one or more of the persons within the data using all the means reasonably likely to be used. In particular, the inquirer might be:

- someone who wished to market cannabis-based medication to a patient who has been prescribed Stiripentol;
- researchers;
- lobbyists and journalists; or
- a family member of a patient who wished to identify the medication being prescribed.

4. Mr Hopkins defined the issue as one of the 'identification risk'. He put his argument in his skeleton argument:

> 26. The dispute is really about … how the test of reasonable likelihood is applied. 'Reasonably likely' is a broad term, capable of spanning a considerable margin of risk. Thresholds such as 'insignificant' or 'sufficiently remote' (which mean the same thing) are important because they indicate what kinds of risk level will meet the test of reasonable likelihood.

He argued that: (a) the First-tier Tribunal had 'wrongly departed from the approach to identification risk set out in the leading cases … and also in guidance'; and (b) failed to apply the test correctly to the facts.

5. Mr Metcalfe disagreed and defined the issue as one of identifiability rather than risk. He argued that the tribunal did not misdirect itself on the test and did not apply the law wrongly to the facts.

C. Two preliminary points

6. Before coming to the legislation and the case law, I can put two points quickly aside.

There is no bright line rule

7. I am sure that public authorities would like a bright line rule that they could apply to decide whether or not information was disclosable under the Freedom of Information Act. Mr Hopkins told me that he was not arguing for that approach. He accepted that the issue had to be decided by reference to the relevant legal concepts and not by reference to a numerical number relating to incidents in the data. I accept that as the correct position in law: see what Cranston J said about the issue not

being a purely statistical one in the *Department of Health* case, which is discussed later in [25] and following.

There is no need to refer to guidance

8. I was referred to guidance, but do not need to refer to it. It is not binding, so it is the law that matters. Moreover, I understand that the Information Commissioner is likely to issue redrafted guidance on the issue in this case, so discussion of the current guidance will not be useful.

D. What the legislation says

Freedom of Information Act 2000

9. Section 40(2) is the relevant exemption in this case. It is a sufficient summary to say that it provides an exemption for personal data that is requested by someone who is not the subject of that data. Mr Hopkins accepted that if this exemption applies, the exemption in section 41 (information provided in confidence) would also apply. He confined his arguments to section 40(2).

Data Protection legislation

10. Section 3 of the Data Protection Act 2018 defines personal data:

> (2) 'Personal data' means any information relating to an identified or identifiable living individual (subject to subsection (14)(c)).
>
> (3) 'Identifiable living individual' means a living individual who can be identified, directly or indirectly, in particular by reference to—
>
> (a) an identifier such as a name, an identification number, location data or an online identifier, or
>
> (b) one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of the individual.

11. This is in line with the definitions in the General Data Protection Regulation (EU) 2016/679. Recital 26 to the Regulation is relevant, because it refers to identifiability and to the means that should be taken into account:

> (26) The principles of data protection should apply to any information concerning an identified or identifiable natural person. Personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person. To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments. The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a

> manner that the data subject is not or no longer identifiable. This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes.

Just looking at the legislation

12. Section 3 of the 2018 Act creates a binary test: can a living individual be identified, directly or indirectly? If the answer is 'yes', the data is personal data. Otherwise, it is not. That is what the Act says, and it is consistent with the Regulation. There is no mention of any test of remoteness or likelihood.

13. The test has to be applied on the basis of all the information that is reasonably likely to be used, including information that would be sought out by a motivated inquirer, as in this case. That derives from Recital 26.

E. What the cases decide

14. Mr Hopkins relied on five cases. This is why they do not support his argument and do not displace what the legislation provides.

*Breyer v Federal Republic of Germany (*Case C-582/14 EU:C:2016:779) [2017] 1 WLR 1569

15. One of the issues for the Court was whether a dynamic IP address registered by an online media services provider when a person connected to a publicly accessible website constituted personal data.

16. The English version of the Court's judgment reads:

> 45. However, it must be determined whether the possibility to combine a dynamic IP address with the additional data held by the internet service provider constitutes a means likely reasonably to be used to identify the data subject.
>
> 46. Thus, as the Advocate General stated essentially in point 68 of his Opinion, that would not be the case if the identification of the data subject was prohibited by law or practically impossible on account of the fact that it requires a disproportionate effort in terms of time, cost and man-power, so that the risk of identification appears in reality to be insignificant.

Despite what the Court said, that is not quite what the Advocate General said:

> 68. Just as recital 26 refers not to any means which may be used by the controller (in this case, the provider of services on the Internet), but only to those that it is likely 'reasonably' to use, the legislature must also be understood as referring to 'third parties' who, *also in a reasonable manner*, may be approached by a controller seeking to obtain additional data for the purpose of identification. This will not occur when contact with those third parties is, in fact, very costly in human and economic terms, or practically impossible or prohibited by law. Otherwise, as noted earlier, it would be virtually impossible to discriminate between the various means, since it would always be possible to imagine the hypothetical contingency of a third party who, no matter how inaccessible to the provider of services on the Internet, could — now or in the future — have additional relevant data to assist in the identification of a user.

17. The difference between the Court and the Advocate General is this. The latter gave three instances of when linking information would not lead to identification – in

summary, they were cost, impracticality and illegality. The Court linked impracticality with cost, linking the latter to the former as cause and effect. At least, that is what the English version says.

18. In other versions that I have checked – French, German, Italian and Spanish – the wording is different. There cost is merely an example of impracticality. Inserting the appropriate words changes [46] to read:

> 46. Thus, as the Advocate General stated essentially in point 68 of his Opinion, that would not be the case if the identification of the data subject was prohibited by law or practically impossible *for example* on account of the fact that it requires a disproportionate effort in terms of time, cost and man-power, so that the risk of identification appears in reality to be insignificant.

I read the paragraph in that way, which also honours the Court's reference back to the Advocate General's opinion.

19. This is the most authoritative of the decisions cited. The Court referred to 'risk of identification' and used the term 'insignificant'. But it identified the controlling factor as the means that were taken to be into account 'to identify the data subject.' I read the judgment and especially paragraph [45] as saying that it must be able to identify the actual data subject by using the means available.

20. There was an argument before me whether the Court was talking about means or outcome. What I take from the judgment is this. Means and outcome are inevitably linked. Speaking of one, inevitably involves speaking of the other. The chance of a particular outcome depends on the means that can be employed and the means available controls the potential outcome. By limiting the means that can be employed, the chances of identification are reduced.

21. That is not, though, the same thing as imposing an additional test of remoteness or significance or likelihood. Eliminating those means will exclude any possibility of identification that is insignificant. Similarly, if this is different, any possibility that is extremely remote is also excluded. But the test remains whether it is possible to identify a specific individual solely by relying on the data available.

22. Identifying a pool that contains or may contain a person covered by the data is not sufficient. Saying that it is reasonably likely that someone is covered by the data is not sufficient. Still less is it sufficient to say that it is reasonably likely that a particular individual may be one of the pool. Linking any specific individual to the data in any of these circumstances does not rely solely on the data disclosed and other data available by reasonable means; it involves speculation. This is the point that the tribunal was making when it referred to guessing. Any break in the chain between the information and the data subject can only be bridged by speculating or guessing. That is especially likely to arise when there is a pool of potential subjects.

Common Services Agency v Scottish Information Commissioner [2008] 4 All ER 851

23. This case presents an initial difficulty on account of the differing views of the judges. For the record, I agree with and adopt to the approach of Cranston J in *R (Department of Health) v Information Commissioner* [2011] EWHC 1430 (Admin) at [45]. In short, he decided that Lord Hope's speech was 'determinative'.

24. Having said that, I find nothing in that case to assist me with this appeal. The issue was the significance of applying a process of barnardisation in order to

anonymise data relating to the incidents of leukaemia in census wards in a particular area. The House of Lords sent the case back to the Commissioner to apply its approach, but did not give any guidance on the issue before me. There may be occasional phrases that can be picked out, but the words were never chosen with my particular issue in mind. As such, they provide no precedent authority.

R (Department of Health) v Information Commissioner [2011] EWHC 1430 (Admin)

25. This is the case that I mentioned (in [23]) and anticipated (in [7]) earlier. It concerned the disclosure of data relating to abortions performed after 24 weeks gestation. Cranston J found that the Information Tribunal had misdirected itself in law, but that the mistake was not material:

> 55. Thus, on this issue, the Tribunal was wrong in its interpretation of the law. As I shall explain shortly, however, it was entitled to arrive at the conclusion that it was extremely remote that the public to whom the statistical data was disclosed would be able to identify individuals from it. In other words: the requested statistics were fully anonymised. It follows that the Tribunal ought to have held that the disclosure of the information to the public did not constitute the processing of personal data.

The judge gave more detail later:

> 70. Moreover, there was no example within the past of identification from published statistical information, nor was there any evidence of information in the public domain that could be used in conjunction with these statistics so as to identify individual patients and doctors. The Tribunal evaluated the Department's argument that published statistics could make a significant contribution to the chain of events, but rejected it in relation to both the Jepson and Nine Year Old Girl cases. It did not misinterpret the figure of 10 for a safe cell, or fail to consider the safety of values of five and below, but took the view that the safety threshold was not dependent on statistical expertise alone. It concluded that the possibility of identification by a third party from these statistics was extremely remote, regardless of the frequency of cell numbers, whether the value was zero to five, or 90 to 100. In summary, there is no legal flaw in its handling of the evidence.

26. The judge provides two examples of relevant evidence: absence of any past identification and absence of any specific additional information that could lead to identification. Both have to be assessed in the context of the evidence as a whole. But it is significant that both examples relate to actual identification of specific individuals.

27. There is no denying that the judge referred to extreme remoteness in both [55] and [70]. I do not accept that the case is an authority for that proposition that a test of remoteness applies in the possibility of identification, thereby qualifying the test set out on the face of the legislation. The judge was certainly not doing that in [55], which is purely descriptive. He was there picking up the language he had used in [32] to summarise the tribunal's reasoning - it is not clear whether those were the tribunal's actual words. His remarks in [70] are, I admit, more difficult to explain away. I do not, though, accept that he was laying down a legal test. If that is what he was doing, he was qualifying the apparently absolute statutory requirement of actual identification

by adding a new element with no basis in previous authority, and doing so without any analysis. This is not how judges lay down or develop the law.

R (Bridges) v Chief Constable of South Wales Police [2020] 1 All ER 864

28. Only the judgment of the Divisional Court is relevant. The issue I have to decide did not arise when the case came before the Court of Appeal at [2021] 2 All ER 1121. The case concerned the use of automated facial recognition. One issue was whether there had been a breach of the Data Protection Act 1998. The Court considered whether the use of the software might be taken together with other information so that the claimant could be identified. The Court cited [40] to [42] and [45] to [49] of *Breyer*, picking up on the final words of paragraph [46]:

> 118. Thus, the only incidents excluded were where the risk of identification 'appears in reality to be insignificant'.

The Court rejected the claimant's argument, saying:

> 123. … the possibility of indirect identification … is somewhat speculative.

29. The Court picked up and applied the language of *Breyer*. It adds nothing to the analysis in that case beyond providing a specific example of the application of the European Court's test.

Information Commissioner v Miller [2018] UKUT 229 (AAC)

30. This case concerned homelessness statistics for the financial years 2009/2010, 2010/2011 and 2011/2012. The request was made in December 2015. The tribunal found that the data was not personal data. Upper Tribunal Judge Markus dismissed the appeal.

31. Interpreting what the judge said takes care. She reminded herself at [26] that she had to exercise 'a certain degree of restraint' in examining the First-tier Tribunal's reasons. She then embarked on an analysis of that reasoning, in the course of which she explained why the language used by the tribunal, when properly understood in its context, showed that the tribunal had not misdirected itself in law.

32. Mr Hopkins relied on this paragraph:

> 28. The FTT's reasons also make it clear that it was aware that the test it had to apply was the risk, on publication of the data, of a member of the public identifying any individual on the basis of that data along with data *other than* that which is in the possession of the data controller. Thus at paragraph 30(d) the FTT set out the Appellant's submission that "it would be impossible for a third party to identify individuals from this data, without knowing very specific details about their circumstance at that point in time, for example the outcome of their application." This shows that the FTT was considering the risk of identification on the basis of the data and other information in the hands of a third party. The DCLG's reasoning which the FTT cited at paragraph 33(b) addressed the likelihood of identification from the data along with other information ("data matching or similar techniques"). Paragraph 34 makes it clear that the FTT was considering whether there was a "reasonable likelihood" that disclosure of the information would result in disclosure of personal data of an identifiable individual, and the penultimate sentence of paragraph 35 shows that the FTT was aware that the question must be considered by reference to the

> data 'together with any other data available'. Asking whether there is a 'reasonable likelihood' of identification may not in all cases be the same as asking whether a person can be identified taking account of 'all means reasonably likely to be used' but in this case it is clear that the First-tier Tribunal was addressing substantively the correct question: what are the chances of an individual being identified? That was in substance the correct focus, particularly given the way in which Cranston J expressed the decision in *Department of Health*: that identification was 'extremely remote' (paragraph 55), and the approach in the Commissioner's Code of Guidance that 'the risk of identification must be greater than remote and reasonably likely' (page 16).

33. I have to read what Judge Markus said in its context, just as she had to read the First-tier Tribunal's reasons. The tribunal had asked whether there was a reasonable likelihood that disclosure would lead to an identifiable individual and Judge Markus found no error in that approach. however, she said that the correct question was:

> whether a person can be identified taking account of 'all means reasonably likely to be used'.

In other words, she was distinguishing between the test to be applied and the way that question might present itself in the context of a particular case. She went on to refer Cranston J's 'extremely remote' test, but I have already explained why that is not part of the test.

F. Mr Hamed's evidence

34. Mr Hamed is a statistician who works for NHSBSA. He has 17 years of experience in a variety of organisations over a range of social and economic sectors, as an analyst, economist and statistician. He made a witness statement, which was filed on behalf of NHSBSA in the First-tier Tribunal. The statement covers two topics.

35. The first topic is a demonstration of the way in which data from other sources could be used in conjunction with the withheld date to product likely matches with real people. As Mr Hamed acknowledges, the matches are 'likely' and not certain, because there could be other explanations for the information available that do not link it to a specific individual. For example, the person identified may not be receiving treatment on the NHS or may be receiving different treatment or receiving medication through a hospital. This evidence was fatal to NHSBSA's case, given that the test requires actual identification. Mr Hamed accepted that it was not possible to identify a specific individual, because there were a variety of possible reasons for the person to appear in the data.

36. The other topic covered by the statement is the common practice used to prevent personal data being disclosed:

> 19. Due to the risks outlined above the risk of identification for data that includes single instances of activities or items that relate to one person is generally considered to be large and it is rare in my experience that official statistical sources would release information without taking measures to reduce the risk. To manage this risk it is commonplace for statistical publications and releases to redact or round small numbers, or otherwise adapt tables and results so that numbers lower than some threshold are usually suppressed.

As I said in [7], Mr Hopkins accepted that this is not what the law requires.

G. Conclusion

37. The legislation provides that actual identification is necessary in order for data to be personal data. The cases cited are consistent with that proposition. They certainly do not require me to decide otherwise. And there is no reason of principle or policy to justify departing from the language of the legislation.

38. The tribunal did not misdirect itself in substance. I do not agree with every word of its analysis of the law, but that does not matter. It came to the correct conclusion that identification of an actual data subject was required. Nor did the tribunal fail to apply that test correctly to the facts. Given Mr Hamed's evidence, it came to the correct conclusion, indeed the only conclusion it could properly reach.

Signed on original
on 06 August 2021

Edward Jacobs
Upper Tribunal Judge