



Public Health
England

Protecting and improving the nation's health

HCAI DATA CAPTURE SYSTEM

User Manual

Find Duplicates Report

Version 1.0

About Public Health England

Public Health England exists to protect and improve the nation's health and wellbeing, and reduce health inequalities. It does this through world-class science, knowledge and intelligence, advocacy, partnerships and the delivery of specialist public health services. PHE is an operationally autonomous executive agency of the Department of Health.

Public Health England
Wellington House
133-155 Waterloo Road
London SE1 8UG
Tel: 020 7654 8000
www.gov.uk/phe
Twitter: @PHE_uk
Facebook: www.facebook.com/PublicHealthEngland

© Crown copyright 2015

You may re-use this information (excluding logos) free of charge in any format or medium, under the terms of the Open Government Licence v2.0. To view this licence, visit [OGL](https://www.nationalarchives.gov.uk/ogl/) or email psi@nationalarchives.gsi.gov.uk. Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.



Document History

Revision date	Author	Version
01/09/2015	Public Health England	1.0
01/10/2015	Public Health England	2.0

Accessing the Duplicates Report

Introduction

The Find Duplicates report is available to all data entry users and other users based on their roles and permissions. The purpose of the Find Duplicates report is to enable users to check whether there are duplicate reports on the HCAI data capture system (DCS) for cases within **their** organisation. For users with permission to enter data for more than one organisation all of this data will be available. Case capture allows potential duplicates to be identified before they are saved, although it is possible to over-ride the duplicate check at this point. This report therefore identifies potential duplicates that have been entered and saved onto the HCAI DCS.

It is necessary to ensure that duplicates, as defined below, are not retained on the HCAI DCS as this will over-inflate the number of cases.

Duplicate Definition

The definition of a duplicate for:

- MRSA, MSSA and *E. coli* bacteraemia is any specimen collected from the same patient within **14 days**.
- CDI a duplicate is any specimen collected from the same patient within **28 days**.

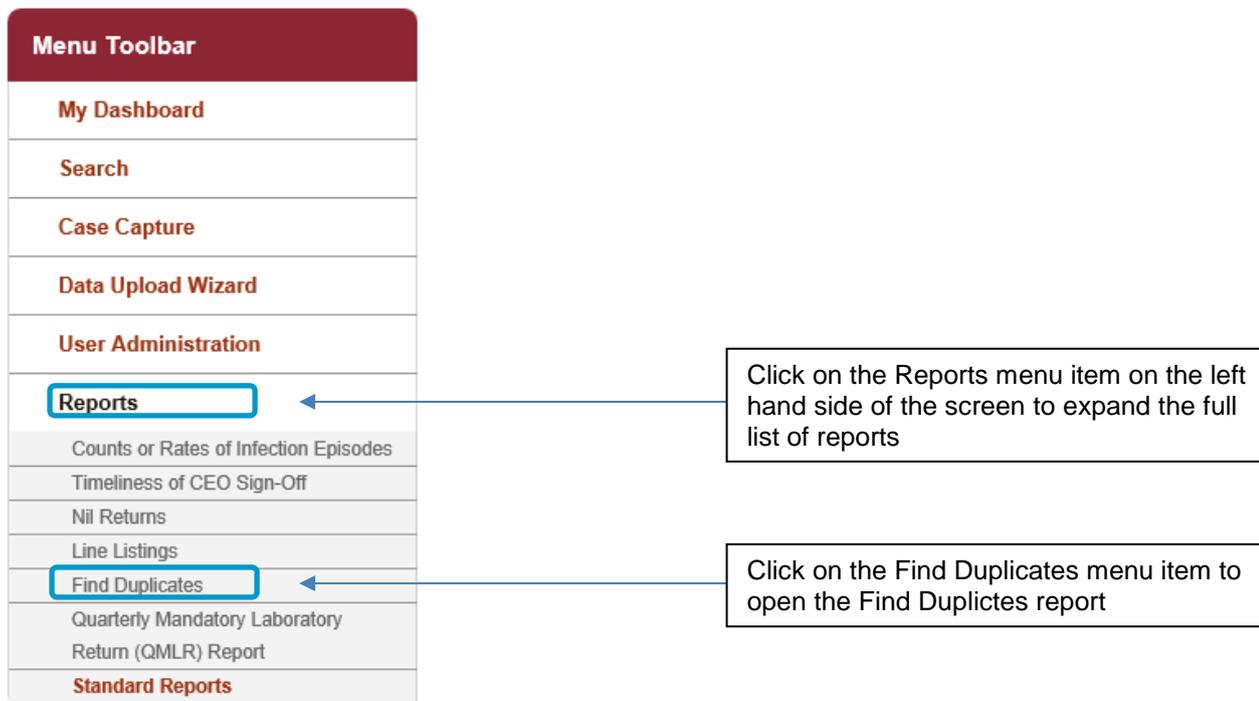
Any positive specimens collected after 14 days (MRSA, MSSA, *E. coli Klebsiella spp. and Pseudomonas aeruginosa* bacteraemia) or 28 days (CDI) are considered to be a new episode and must be entered onto the HCAI DCS.

Using the Find Duplicates report

To access the “Find Duplicates” report, from the home screen click on ‘Reports’ in the ‘Menu Toolbar’, then click on ‘Find Duplicates’ (Figure 1.1).

Figure 1.1: Accessing the Duplicate Report

Finding Duplicates Report (Figure 1.2)



The report allows you to search for duplicates by a number of different permutations, including organisation, specimen date and sector (see Table 1.1) and as well as the patient and specimen-specific identification criteria (see Table 1.2)

When you first land on the Finding Duplicates report page, a report based on user-specific pre-set parameters, similar to Figure 1.2 will appear. Table 1.1 provides an overview of the report parameters and their impact on what data are presented.

Please note that one or more of the parameters can be selected and they work as additive – ie selecting two parameters will return cases matching both of those criteria. Any of the reports returned after selecting the identification criteria have been selected also need to fulfil the criteria in the report parameters.

Also state at the start that records are returned that are within the episode length

Figure 1.2: Finding Duplicates report screen

The screenshot shows a web-based form titled "Find Duplicates". It contains several input fields and a button:

- Period From:** 13/10/2014
- Period To:** 13/10/2015
- Organisation Type:** NHS Trust
- Region:** LONDON
- Organisation:** BARKING, HAVERING AND REDBRIDGE
- Data Collection:** MRSA
- Sector:** NHS
- Identification Criteria:** NHS Number, Date of Birth, Patient Name
- View Report:** A blue button on the right side of the form.

Table 1.1: Finding Duplicates: Report Parameters

Parameter	Comment	Parameter Type	Note
Period from / Period to	Changes the time period the report displays the data	Date field or drop down date box as in Figure 1.3	It is possible to search for a specific date or date range. By default a 12 month period is selected. NB if specimen date isn't selected in the identification criteria then duplicates outside the time range won't be identified. If specimen date isn't selected then episode length isn't considered in the criteria
Organisation Type	Changes the organisation type the report displays the data by.	Drop-down	The default value is the organisation type that you are currently logged into the system as.
Region	Changes the region that the data are presented for in this report. These include the NHS Regions and an overall 'National' value.	Multi-select drop-down	The default value is the region associated with the organisation you are signed in as. Depending on user roles and permissions it may be possible to selected other regions.
Organisation	Changes the organisation for	Multi-select drop-down	The default organisation will be the

	which this report presents data.	NB. The drop-down menu can be made larger, so that you can read the full name of the organisations, by clicking on and dragging the bottom right-hand corner to a wider point. 	organisation that you are signed in as. If you have access to multiple organisations of the same type, then all of the organisations will be selected as the default. The organisation drop-down options are based on the organisation type and region that you have already selected.
Data Collection	Changes which Data Collection the data is presented for in this report.	Drop-down	Options are MRSA, MSSA, <i>E.coli</i> , <i>Klebsiella spp.</i> , <i>Pseudomonas aeruginosa</i> & <i>C. difficile</i>
Sector	Changes the sector for which this report presents data.	Drop-down	The default value will depend on the 'Organisation Type' that you have already selected; if the organisation type is 'Independent Sector Healthcare Provider' or 'Independent Sector Healthcare Provider Site' then the default Sector will be 'Independent'. For all other 'Organisation Types' selected, the default sector is 'NHS'
Identification Criteria	Allows you to select which fields you would like to be included in the report.	Multi-select drop-down NB. The drop-down menu can be made larger, so that you can read the full name of the fields, by clicking on and dragging the bottom	The default value is '(Select All)' Please refer to Table 1.2 for further detail

		right-hand corner to a wider point. 	
--	--	---	--

Table 1.2 Finding Duplicates – Identification Criteria

Parameter	Comment
NHS Number	Selecting NHS number will include reports within the episode period where the same NHS number has been entered; this included reports with an NHS number of all 9s.
Date of Birth	The patient's date of birth
Patient Surname Soundex	The soundex is an encrypted version of the patients surname, taking the form of a letter followed by 3 numbers. Please note that as more than one surname can produce the same soundex this is not a unique identifier.
Specimen Date	The specimen date around which you wish to search for duplicate cases. Duplicates within the episode length of this specimen date will be returned.
Specimen Number	The specimen number relating to the case entered onto the HCAI DCS.
Hospital Number	The patient's hospital number. A local identifier.

Figure 1.3: Finding Duplicates: Screen View

Find Duplicates

Period From: 07/10/2014 Period To: 07/10/2015 [View Report](#)
 Organisation Type: NHS Trust Region: NATIONAL
 Organisation: AINTREE UNIVERSITY HOSPITAL Data Collection: C. difficile
 Sector: NHS Identification Criteria: NHS Number, Date of Birth, P

1 of 1 Find | Next

Public Health England

Find Duplicates

Period - From	07/10/2014	Period - To	07/10/2015	Region	NATIONAL
Data Collection	C. difficile	Organisation Type	NHS Trust	Sector	NHS
Organisation	-All-	Identification Criteria	NHS Number Date of Birth Patient Surname Soundex Specimen Date Specimen Number Hospital Number		

ID	Reporting Organisation	Region	Clinical Commissioning Group	Data Collection	NHS Number	Patient Name	DOB	Age	Age Description	Data Collection Date
39	BARNSELY HOSPITAL NHS FOUNDATION TRUST	NORTH OF ENGLAND	NHS ROTHERHAM CCG	C. difficile		J S520		87	87 years	20/01/2015
38	THE ROTHERHAM NHS FOUNDATION TRUST	NORTH OF ENGLAND	NHS ROTHERHAM CCG	C. difficile		J S520		87	87 years	20/01/2015

Figure 1.6: Finding Duplicates: Export Icons

Search bar & controls Export icon → when you click on this, the following options appear:

Arrows to flick between pages of the report, if there are more than one.

Allows you to change the size of the presented report, but does not change the number of observations shown on screen

Refresh icon

Assessing whether two records are Duplicates using the Identification Criteria

The report output will return cases matching the criteria selected. A number of fields are presented in the output to help you review in more detail evidence for duplicates. It is possible to drill down to the case capture screen of individual records by clicking on the hyperlinked Case ID. This allows further patient level information to be viewed to determine evidence for duplication.

Figure 1.4: Finding Duplicates output

Find Duplicates

Period From

Organisation Type

Organisation

Sector

Period To

Region

Data Collection

Identification Criteria

[View Report](#)

1 of 1
Find | Next

Find Duplicates

Period - From	07/10/2014	Period - To	07/10/2015	Region	NATIONAL
Data Collection	C. difficile	Organisation Type	NHS Trust	Sector	NHS
Organisation	-All-	Identification Criteria	NHS Number Date of Birth Patient Surname Soundex Specimen Date Specimen Number Hospital Number		

ID	Reporting Organisation	Region	Clinical Commissioning Group	Data Collection	NHS Number	Patient Name	DOB	Age	Age Description	Data Collection Date
39	BARNSELEY HOSPITAL NHS FOUNDATION TRUST	NORTH OF ENGLAND	NHS ROTHERHAM CCG	C. difficile	[REDACTED]	J S520	[REDACTED]	87	87 years	20/01/2015
38	THE ROTHERHAM NHS FOUNDATION TRUST	NORTH OF ENGLAND	NHS ROTHERHAM CCG	C. difficile	[REDACTED]	J S520	[REDACTED]	87	87 years	20/01/2015

Please note: The details in this example have been obscured as the data were taken from the live CDI dataset (NB: episode length is 28 days).

When a Duplicate is identified

A duplicate record can be deleted by users with the relevant permissions if the record is not part of a locked dataset (i.e. a period that has been signed off).

Delete a case.

Using the 6 digit unique ID, search for the case to be deleted using the search option in the menu toolbar.

	Condition	Data Collection	ID	Data Collection Date	Region	
<input type="checkbox"/>	[All]	[All]	[All]	[All]	[All]	[All]
<input checked="" type="checkbox"/>	Staphylococcus aureus	MRSA	[REDACTED]	17-Sep-2017	LONDON	[REDACTED]

View [REDACTED]

Page: 1 of 1 Go Page size: 1 Change Item 1 to 1 of 1

Cancel **Delete** New Infection Episode

However, if the record which needs to be deleted is part of a locked dataset users must contact mandatory.surveillance@phe.gov.uk to request the dataset to be unlocked. Duplicates identified soon after "Sign Off" may be removed, however if the duplicate has been undetected for greater than 6 months then the individual circumstances will need to be considered.

Please note: that once a report is deleted the process cannot be undone and if deleted in error the case will need to be entered from scratch.