



# Cervical screening: invasive cervical cancer audit 2013 to 2016

## Appendix A: data completeness

### Data completeness and limitations

When considering the findings presented in this report, the varying degree of completeness of the available information should be considered. The difficulties involved in ensuring the completeness of essential data fields are described below.

It is rare for data to be reported as missing, but missing data should be distinguished from incompleteness of record. Missing data may be unavailable (for example where a death certificate which does not provide information about cancer staging has been used), or may not yet have been recorded as part of the audit. For this reason, we have used the term 'none recorded' to describe cases where final stage is still pending and the term 'none available' to describe cases where after considerable effort, no staging data has been available.

Other cases may be subject to reporting delays, having been submitted to the audit before all essential fields could be completed. In these instances, missing fields are updated as and when data become available, with the result that complete information may not be received for some months after the case has been registered. An additional challenge, which can create further delay, is the need to coordinate between the various aspects of the audit process when a case of cervical cancer is diagnosed.

### Dealing with missing values

Cases reported in the MB1 series (Cancer Registration Statistics in England, Office for National Statistics) between 2013 and 2015 were compared to those recorded in the audit for the same period, by age at diagnosis (Table A-1a). The aim was to ascertain whether there is a subset of cases for which a delay in their inclusion in the audit is more likely, and whether this is related to age at diagnosis. 94% of the cervical cancer

cases registered in England between 2013 and 2015 were recorded in the audit for the same period. However, the audit data are more likely to include cases diagnosed in women between the ages of 25 and 64 (audit includes 98% of all registered cancers in this age group), than cases diagnosed in women over the age of 65 (audit includes 75% of registered cancers in this age group). The completeness of the audit data, compared with MB1 decreased with increasing age at diagnosis.

We assessed the completeness of audit data for FIGO stage by comparing the distribution of staged cancers diagnosed between April 2011 and March 2012 across 4 audit years (Table A-1b). The table shows that if we were to assign a stage to cases with this information missing, assuming that stage was missing at random, we would be overestimating the proportion of cases diagnosed with early stage cancer. For example, based on data received as of October 2011 we would have assumed 47.6% of cases diagnosed between April 2011 and March 2012 had stage IA cancer. However, by October 2012 that proportion decreased to 42% and remained as such since then. This suggests that cases with unknown FIGO stage are more likely to be advanced stage cancer.

In the audit reports published in July 2011 and May 2012, we assumed that data for FIGO staging was missing at random, which would have led to an overestimation of the proportion of stage IA cancers and an underestimation of the proportion of stage II+ cancers. For this report (and the 2 previous reports), we have used a more complicated model that takes into account the differential delays in obtaining stage.<sup>1</sup>

---

<sup>1</sup> A multinomial logistic regression model was fitted with outcome 'stage at diagnosis' and explanatory variables age group, treatment type and year of diagnosis. Using the results of this model, the probability of each stage category was then predicted for each individual with missing stage.

**A-1a National Cancer Registration Statistics (MB1 Series), published by the Office for National Statistics, compared with those reported in the audit**

Total cases reported*	Age at diagnosis															Total	25-64
	<20	20-24	25-29	30-34	35-39	40-44	45-49	50-55	55-59	60-64	65-69	70-74	75-79	80-84	85+		
MB1 series 2013 to 2015 (n)	9	221	1208	999	901	842	707	540	456	387	359	288	321	258	250	7746	6040
Audit 2013 to 2015 (n)	3	216	1238	1016	903	814	695	509	420	350	309	211	243	183	160	7270	5945
Difference (n)	6	5	-30	-17	-2	28	12	31	36	37	50	77	78	75	90	476	95
Proportion (%)	33.3	97.7	102.5	101.7	100.2	96.7	98.3	94.3	92.1	90.4	86.1	73.3	75.7	70.9	64.0	93.9	98.4

\* MB1 Cancer Statistics are published by calendar year, audit data are normally reported by financial year (1 April to 31 March).

**A-1b Cancers in women aged 25 to 64, diagnosed between April 2011 and March 2012**

Received as of	Observed stage by year of audit data (n)							Proportion assuming missing at random (%)		
	IA	IB	II	III+	IB+	None recorded	Total	IA	IB	II+
Oct 2011	285	207	50	28	29	124	723	47.6	38.1	14.3
Oct 2012	656	520	169	124	73	161	1703	42.5	36.8	20.7
Oct 2013	691	562	194	152	73	135	1807	41.3	36.3	22.4
Oct 2017	774	601	214	172	71	117	1949	42.2	35.2	22.6

## A-2a Proportion of essential data collected for cases in section A: personal and cancer details

	Section A: Essential fields								
	Cases	Date of Birth		Date of Diagnosis		Stage*		Histology	
		n	%	n	%	n	%	n	%
Current report (Apr 2013 to Mar 2016)	<b>6,028</b>	6,028	100	6,028	100	5,718	94.9	5,490	91.1
Fourth report (Apr 2009 to Mar 2013)	<b>8,784</b>	8,784	100	8,784	100	8,014	91.2	8,543	97.3
Third report (Apr 2009 to Mar 2012)	<b>6,508</b>	6,508	100	6,508	100	5,901	90.7	6,336	97.4
Second report (Apr 2007 to Mar 2011)	<b>8,566</b>	8,566	100	8,566	100	7,423	86.7	8,197	95.7
First report (Apr 2007 to Mar 2010)	<b>6231</b>	6231	100	6231	100	5197	83.4	5922	95.0

\*Cases where data collection is complete, and stage is missing are staged as a reasonable amount of effort has been made to collect the data. Incomplete cases with a stage recorded as X (or missing) are considered not to have stage. Please refer to section 6 for full details regarding missing data.

## A-2b Proportion of data collected for cases in section A: personal and cancer details

Section A: Non-essential fields										
	Treatment (in those with known treatment, excluding those reported as none*)		Treatment (in those with treatment recorded including those recorded as none)		Index of Multiple Deprivation (Cases)		Index of Multiple Deprivation (Controls)			
	Cases	n	%	n	%	n	%	Controls	n	%
Current report (Apr 2013 to Mar 2016)	6,028	3,990	66.2	4,254	70.6	5,224	86.7	11,580	8,451	73.0
Fourth report (Apr 2009 to Mar 2013)	8784	5970	68.0	6183	70.4	6843	77.9	17270	7345	42.5
Third report (Apr 2009 to Mar 2012)	6508	4146	63.7	4394	67.5	5104	78.4	12841	4423	34.4
Second report (Apr 2007 to Mar 2011)	8,566	5,199	60.7	5,675	66.3	6,485	75.7	16,920	7,964	47.1
First report (Apr 2007 to Mar 2010)	6231	3086	49.5	3382	54.3	4723	75.8	12335	5947	48.2

\* Where treatment was recorded as 'None' we assume it means 'none other than palliative care'. Attempts have been made to clarify this issue and there is now a category for palliative care; however, some misclassification may remain and therefore they are excluded from this column.

### A-3 Proportion of cases with FIGO stage reported as none recorded, none available\* or IB or worse (1B+), by age and audit year (from April 2013 to March 2016)

	<b>None recorded</b>	<b>None available</b>	<b>IB+ (NOS)</b>	<b>Total</b>
<b>Age</b>	<b>%</b>	<b>%</b>	<b>%</b>	<b>%</b>
<25	0.0	5.3	3.5	8.8
25 to 49	1.5	4.0	1.7	7.2
50 to 64	2.5	6.6	2.8	11.9
65+	4.9	8.4	3.6	16.8
<b>Audit Year</b>				
2013/14	2.4	3.6	2.2	8.2
2014/15	2.3	4.7	2.5	9.6
2015/16	1.7	7.4	1.8	10.8
<b>Previous reports</b>				
Current report (Apr 2013 to Mar 2016)	2.2	5.1	2.2	9.5
Fourth report (Apr 2009 to Mar 2013)	8.8	1.9	3.8	14.5
Third report (Apr 2009 to Mar 2012)	9.3	1.1	4.4	14.8
Second report (Apr 2007 to Mar 2011)	12.0	1.6	4.2	17.8
First report (Apr 2007 to Mar 2010)	16.6	N/A	4.4	21.0

\* Where stage is reported as none available instead of none recorded a reasonable amount of effort has been made to find the stage, but none has been available. This is derived from cases recorded as 'audit complete' which means that no further details are being sought for these women. The option to report cases as 'none available' has only been available to all SQAS since April 2012.

## A-4 Proportion of data collected for cases in section B: cytology

<b>Section B: cytology</b>								
<b>Completeness of data among recorded cytology tests</b>								
<b>Audit year</b>	<b>Cases</b>	<b>Tests on all cases<sup>a</sup></b>	<b>Date test was taken</b>		<b>Result of Test<sup>b</sup></b>		<b>Action Code<sup>b</sup></b>	
			<b>n</b>	<b>%</b>	<b>n</b>	<b>%</b>	<b>n</b>	<b>%</b>
Current report (Apr 2013 to Mar 2016)	6,028	21,764	21,764	100	21,707	99.7	21,746	99.9
Fourth report (Apr 2009 to Mar 2013)	8,784	35,810	35,810	100	35,803	100	35,781	99.9
Third report (Apr 2009 to Mar 2012)	6,508	26,619	26,619	100	26,619	100	26,594	99.9
Second report (Apr 2007 to Mar 2011)	8,567	34,910	34,910	100	34,910	100	34,870	99.9
First report (Apr 2007 to Mar 2010)	6,231	25,972	25,972	100	25,954	100	25,951	99.9

a Cytology tests known to the audit and taken before diagnosis

b Cytology data obtained directly from 'Exeter' call and recall IT system should have all 3 data fields complete. Missing data, we believe, is the result of inclusion into the audit of cytology tests taken before the programme started in 1988 and a few slides that were found in the laboratory, but not recorded on Exeter. These tests will not have 'action code' as Exeter generates this field.

## A-5 Proportion of data collected for cases in section C: colposcopy

Section C: Colposcopy										
Audit report	Cases with an Action Code of 'suspend'	Cases with a 'suspend' and a colposcopy		Cases with a colposcopy but no suspend n	No. of Colp appts n	Date of colp		Satisfactory exam or DNA*		Colp procedure n
		n	%			n	%	n	%	
Current report (Apr 2013 to Mar 2016)	3,674	2,397	65.2	627	4,378	4,378	100	4,378	100	3,601
Fourth report (Apr 2009 to Mar 2013)	6,073	3,963	65.3	494	6,823	6,823	100	6,823	100	5,479
Third report (Apr 2009 to Mar 2012)	4,523	2,843	62.9	430	5,195	5,195	100	4,347	84	4,287
Second report (Apr 2007 to Mar 2011)	5,884	3,604	61.3	647	7,167	7,167	100	5,942	83	5,620
First report (Apr 2007 to Mar 2010)	4,308	2,412	56.0	557	4,348	4,348	100	3,445	79	3,249

\* DNA, did not attend