
Documentation submitted by NHS Blood and Transplant to Professor Sir Gordon Duff's Independent Review

Records of data analyses conducted by NHSBT

ODR Organ Preferences Data Analysis – Stage 1 (14 April 2010)

ODR Organ Preferences Data Analysis – Stage 2 (16 April 2010)

ODR Organ Preferences Data Analysis – Stage 3 (19 April 2010)

ODR Organ Preferences Data Analysis – Stage 4 (3 May 2010)

SUI web traffic and ODR registrations analysis

10 April - 1 May

10 April - 18 April

27 March - 14 April

Identification of change in DVLA registration pattern

ODR Organ Preferences DVLA Data Issue

***Approach & Assumptions
to the Analysis of Impacted Records on the ODR
Stage 1
DE1699-N1***

14 April 2010

V1.0

Table of Contents

1. Summary of changes	3
2. Reviews and Approvals.....	3
3. Introduction.....	4
4. Purpose of Document.....	4
5. Data Analysis Approach	4
6. Stage 1 – Defining Simple and Safe Assumptions to Ring Fence Potentially Impacted Records.....	5

1. Summary of changes

This section records the history of significant changes to this document. Only the most significant changes are described here.

Version	Date	Author	Description of change
0.1	13/04/10		Draft for discussion with 3 approvers
1.0	14/04/10		Version for sign-off

Where significant changes are made to this document, the version number will be incremented by 1.0.

Where changes are made for clarity and reading ease only and no change is made to the meaning or intention of this document, the version number will be increased by 0.1.

2. Reviews and Approvals

Approved by:

Name	Role	Signature	Date
David Shute	Assistant Director – Transplant Support Services		
Alan McDermott	Director of Business Transformation Services		
Ian Bateman	Assistant Director of Quality		

3. Introduction

An investigation of the organ preferences problem by the IT department resulted in a Problem Statement being documented, which records the coding error that has led to incorrect organ preferences records being loaded onto the ODR. In summary, the data problem is:

- For DVLA(GB) records received since 1999, where specific organ preferences have been requested, the preferences for the following organs have been transposed:
 - a. The registrants Liver preference has been recorded as their Lungs preference
 - b. The registrants Lungs preference has been recorded as their Cornea preference
 - c. The registrants Heart preference has been recorded as their Liver preference
 - d. The registrants Cornea preference has been recorded as their Heart preference

4. Purpose of Document

This document summarises the approach to analysing the ODR records that have been impacted by the fault in processing DVLA(GB)'s organ preference feed. A key part of this approach involves making assumptions about the data on the ODR, and validating those assumptions when necessary.

It has been agreed that the approach needs to be agreed by 3 key people:

1. David Shute on behalf of the ODR business
2. Alan McDermott as IT Director
3. Ian Bateman in his role of Quality Assurance

This document has been drafted to document the approach and facilitate agreement and sign-off of the approach.

5. Data Analysis Approach

The analysis approach that is being adopted to determine which current records on the ODR are impacted by this problem is as follows:

1. Determine some simple and safe assumptions that can be used to ring fence the minority of records that **could** have been impacted by the problem
2. Derive a more sophisticated set of assumptions that can be used to determine whether the ring fenced records are incorrect and require remediation, or in fact are actually correct and should not be remediated – these will be documented in a Stage 2 version of the ODR Organ Preferences Data Analysis
3. Validate the more sophisticated assumptions, as they could be susceptible to error
4. Accept that there will always be a set of records for which we cannot determine whether the organ preferences on the ODR are correct or not

Stage 1 of this document only sets out the simple and safe assumptions mentioned in bullet 1, to enable all approvers to confirm these assumptions and confirm that no further validation is required. Documents covering other stages will record the more sophisticated assumptions as they are validated.

6. Stage 1 – Defining Simple and Safe Assumptions to Ring Fence Potentially Impacted Records

The following key assumptions are being made to eliminate any record that could not be impacted by the data problem:

1. Only specific organ preferences were incorrectly transposed, therefore if the preference of “all” organs is recorded on the ODR, the registrants record will be unaffected by the DVLA data transposition error.
2. Only 4 organ preferences (Heart, Lungs, Liver & Cornea) have been incorrectly transposed. If the registrants preference for all of these organs is recorded as the same value (i.e. all 4 are “yes, or all 4 are “no” or “no preference”) then transposing the flags incorrectly will have no impact on the preferences recorded. Such records can therefore be considered as correct.
3. A registrant can only become a future donor if they are alive and have an active record on the ODR. So any record where the status is not set to alive (e.g. deceased, withdrawn consent, duplicates, etc.) can be excluded from the analysis
4. DVLA(GB) always send DVLA numbers in their data feeds. If a DVLA(GB) feed has ever been received for a registrant, a DVLA number will be present on the record, so any record without a DVLA number can be considered as correct.

The total number of records for live registrants picked up from the ODR at the end of March is 17,080,638. The above assumptions have then been applied to filter out records that can be considered as correct. This leaves **997,014** records that **could** be impacted by the data transposition error. These records are being ring fenced and flagged before the NCC start to handle enquiries and updates to organ preferences from registrants.

At a meeting on 14 April 2010, 1pm, all 3 approvers signed up to these assumptions, giving approval to implement the above assumptions in the Production ODR.

ODR Organ Preferences DVLA Data Issue

***Approach & Assumptions
to the Analysis of Impacted Records on the ODR
Stage 2
DE1699-N2***

16 April 2010

V1.1

Table of Contents

1. Summary of changes	3
2. Reviews and Approvals.....	3
3. Introduction.....	4
4. Purpose of Document.....	4
5. Data Analysis Approach for Stage 2	4
6. Stage 2 – Determining a Set of Sophisticated Assumptions to Derive Records for Remediation	5
7. Validation of Assumptions.....	6

1. Summary of changes

This section records the history of significant changes to this document. Only the most significant changes are described here.

Version	Date	Author	Description of change
0.1	14/04/10		Draft for discussion with 3 approvers
1.0	16/04/10		Version for sign-off
1.1	16/04/10		New version for sign-off – approach to remediation

Where significant changes are made to this document, the version number will be incremented by 1.0.

Where changes are made for clarity and reading ease only and no change is made to the meaning or intention of this document, the version number will be increased by 0.1.

2. Reviews and Approvals

Approved by:

Name	Role	Signature	Date
David Shute	Assistant Director – Transplant Support Services		
Alan McDermott	Director of Business Transformation Services		
Ian Bateman	Assistant Director of Quality		

3. Introduction

An investigation of the organ preferences problem by the IT department resulted in a Problem Statement being documented, which records the coding error that has led to incorrect organ preferences records being loaded onto the ODR. In summary, the data problem is:

- For DVLA(GB) records received since 1999, where specific organ preferences have been requested, the preferences for the following organs have been transposed:
 - a. The registrants Liver preference has been recorded as their Lungs preference
 - b. The registrants Lungs preference has been recorded as their Cornea preference
 - c. The registrants Heart preference has been recorded as their Liver preference
 - d. The registrants Cornea preference has been recorded as their Heart preference

The data analysis is being conducted in a series of stages – this document follows on from Stage 1 of the Data Analysis, which ring-fenced 997,014 records as being potentially impacted by the error.

4. Purpose of Document

This document summarises the approach to Stage 2 of the analysis of the ODR records that have been impacted by the fault in processing DVLA(GB)'s organ preference feed. Stage 2 involves analysing the 997,014 records and determining a subset of these that have definitely been impacted by the error, and require remediation.

The approach needs to be agreed by 3 key people:

1. David Shute on behalf of the ODR business
2. Alan McDermott as IT Director
3. Ian Bateman in his role of Quality Assurance

This document has been drafted to document the approach and facilitate agreement and sign-off of the approach.

5. Data Analysis Approach for Stage 2

The analysis approach that is being adopted to determine which current records on the ODR are impacted by this problem is as follows:

1. Determine a set of sophisticated assumptions that can be used to derive a subset of the ring fenced records that are incorrect and require remediation
2. Validate these assumptions against archived file information if it is available

6. Stage 2 – Determining a Set of Sophisticated Assumptions to Derive Records for Remediation

The following key assumptions about how the system works are being made to determine records that have been impacted by the data problem:

1. The DVLA(GB) transposition problem always occurs on records processed after 01.03.99

Evidence for assumption: Analysis of the database shows that a step change in the percentage of registrants agreeing to donate eyes occurred in early 1999. There is no business reason for expecting a step change at this time. In addition, there was an upgrade to the database platform implemented in early 1999 which required a rewrite of the DVLA processing scripts. Finally, a sample (c.50) of records originating from 2000 onwards were checked with the DVLA after the problem was discovered, and all displayed the transposition problem.

2. If the date of the original source details is the same as the date of the latest change, only one feed for that registrant has been uploaded to the ODR. If this source is from the DVLA (GB), and the source date is after 01.03.99 then the data will be incorrect

Evidence for assumption: Analysis of the DVLA(GB) records has shown that organ preferences are always sent when the DVLA(GB) issue their initial registration record (type 1 record).

3. Any records with a registration date after May 2002 (when the audit trail was introduced) will have an audit trail of updates being made (although this does not include details about updates to organ preferences). Assumptions 5 and 6 can only be validated if this assumption is accepted.

Evidence for assumption: A partial audit trail was implemented which provides a log of when changes have been applied, the source of that change, and other information about what data was changed. Examination of the audit trail logs and the audit trail code however demonstrated that no auditing of changes to the organ preferences was implemented.

4. The audit trail records the date of the update, plus the date of the previous update. If the date of the previous update for a registrant's first audit record is the same as the registration date, there is a complete audit trail that includes all updates for that registrant

Evidence for assumption: The audit trail records the value of the new data, and also the old value that is being superseded. As the audit trail logs changes to the latest change date, it is possible to determine whether a full audit trail exists even when a record was registered before May 2002 providing that the first audit trail points to the previous update as being the initial registration.

5. If the latest source is DVLA (GB) after 01.03.99, and no earlier records are from the DVLA (GB), then the latest source is a DVLA (GB) initial registration, so the data will be incorrect

Evidence for assumption: DVLA (GB) have confirmed that they always send a full "type 1" initial registration record (which includes organ preferences) when they first send a record on a registrant to NHS BT.

6. If the latest source is DVLA (GB), or a source that does not update specific organ preferences, and earlier records are only from the DVLA (GB) or other sources that do not update organ preferences, then the originating DVLA (GB) source cannot have been corrected by later sources, so the data will be incorrect

Evidence for assumption: Some sources never send specific organ preferences. These sources therefore cannot have corrected any DVLA (GB) data that has been incorrectly applied

Out of the 997,014 records ring fenced during Stage 1, a total of 523,373 records were identified for remediation.

7. Validation of Assumptions

The above assumptions were drawn up based on a system that is largely undocumented, has incurred a data issue that has been undiscovered for 11 years, and where the information that is available (audit trails, change history) can at times be patchy and difficult to draw conclusions from. In addition, even well documented systems have a tendency to incur data inconsistencies which cannot easily be explained by the experts that support them. So although these assumptions seem reasonable and no-one has contradicted them, they needed to be validated.

Initially, the validation approach was based on sampling small numbers of records compared against archive records to prove that the assumptions were correct. However, discussions with the Statistics department concluded that unless large numbers of records were checked, there was still a danger of an unacceptable level of erroneous updates being applied when rolled out across the full 523,373 selected for remediation.

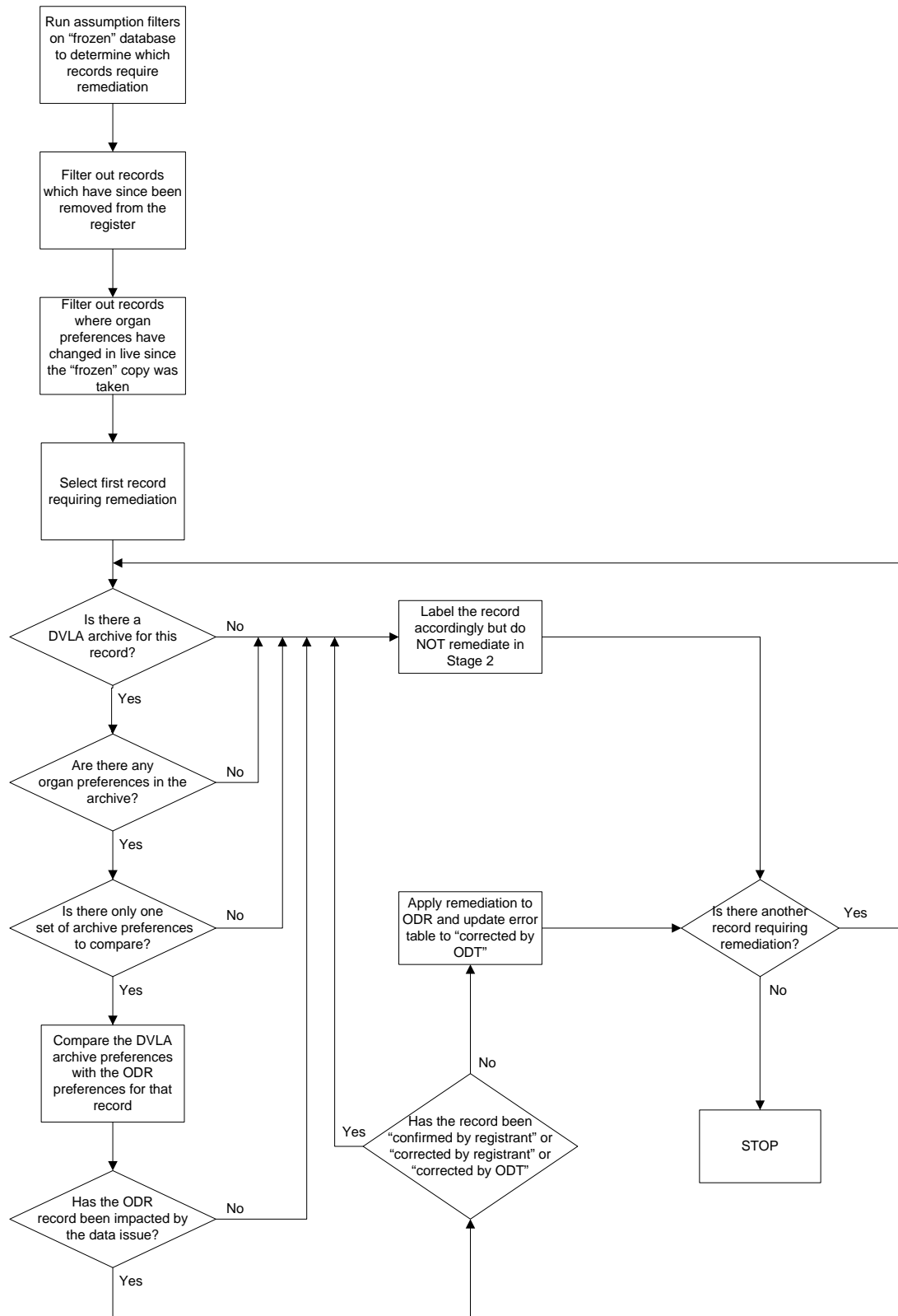
So a different approach to validation was then adopted. Where original source archive is available, this information is compared directly with the ODR database. In all cases selected for remediation, we expect to see that the ODR record displays the transposition error when compared to the source archive record.

An initial run of 288,892 records revealed matching archive records for 228,075 (79%) of that sample. Of the 228,075, 226,887 records (99.5% of archive matches) displayed the transposition error, so were correct to be selected for correction. However, 927 would have been updated incorrectly (the remaining 261 are inconclusive as multiple archive records were found), suggesting a success rate of 99.5%-99.6% if applied to other records where archive file information is not available. The 927 errors includes a list of unaudited ODT transactions, some of which are captured on a paper list and are being converted to electronic format. Once this is done, a lower error rate is expected.

Figures for the remaining 234,481 records (which are based on more complex system assumptions) revealed that DVLA archives were available for 191,703 (82%). Of these 191,703, 165,956 (86.6%) matched exactly. There were 4,047 confirmed errors (2.1%) and a larger number (21,700) of inconclusive records which could be analysed further if more time is available.

The approach for Stage 2 therefore is to remediate only records where an archive record can be found, and update only those where the transposition error is definitely occurring. This means remediating 392,843 records (226,887 plus 165,956). Other records will be left unchanged, and continue to be flagged as “unknown”, and passed through to Stage 3 for consideration.

By signing off this document, all approvers are agreeing to the remediation of the 392,843 records based on the logic defined in this document. The process that will be followed for this remediation is covered off in other documentation, but a high level process flow is attached below for reference



ODR Organ Preferences DVLA Data Issue***Approach & Assumptions
to the Analysis of Impacted Records on the ODR
Stage 3
DE1699-N3******19 April 2010******V1.0***

Table of Contents

1. Summary of changes	3
2. Reviews and Approvals.....	3
3. Introduction.....	4
4. Purpose of Document.....	4
5. Data Analysis Approach for Stage 3	4
6. Key Assumptions for Remediation of Stage 3 Records	5
7. Validation of Assumptions.....	5

1. Summary of changes

This section records the history of significant changes to this document. Only the most significant changes are described here.

Version	Date	Author	Description of change
0.1	19/04/10		Draft for discussion with 3 approvers
1.0	19/04/10		Version for sign off

Where significant changes are made to this document, the version number will be incremented by 1.0.

Where changes are made for clarity and reading ease only and no change is made to the meaning or intention of this document, the version number will be increased by 0.1.

2. Reviews and Approvals

Approved by:

Name	Role	Signature	Date
David Shute	Assistant Director – Transplant Support Services		
Alan McDermott	Director of Business Transformation Services		
Ian Bateman	Assistant Director of Quality		

3. Introduction

An investigation of the organ preferences problem by the IT department resulted in a Problem Statement being documented, which records the coding error that has led to incorrect organ preferences records being loaded onto the ODR. In summary, the data problem is:

- For DVLA(GB) records received since 1999, where specific organ preferences have been requested, the preferences for the following organs have been transposed:
 - a. The registrants Liver preference has been recorded as their Lungs preference
 - b. The registrants Lungs preference has been recorded as their Cornea preference
 - c. The registrants Heart preference has been recorded as their Liver preference
 - d. The registrants Cornea preference has been recorded as their Heart preference

The data analysis is being conducted in a series of stages – this document follows on from Stage 1 of the Data Analysis, which ring-fenced 997,014 records as being potentially impacted by the error, and Stage 2, which remediated 392,843 records which were validated as definitely incorrect.

4. Purpose of Document

This document summarises the approach to Stage 3 of the analysis of the ODR records that have been impacted by the fault in processing DVLA(GB)'s organ preference feed. A key part of this approach involves making assumptions about the data on the ODR, and validating those assumptions when necessary.

The approach needs to be agreed by 3 key people:

1. David Shute on behalf of the ODR business
2. Alan McDermott as IT Director
3. Ian Bateman in his role of Quality Assurance

This document has been drafted to document the approach and facilitate agreement and sign-off of the approach.

5. Data Analysis Approach for Stage 3

The analysis approach that was adopted for Stage 2 to determine ODR records impacted by the data problem was:

1. Determine a set of sophisticated system assumptions that can be used to derive a subset of the ring fenced records that are incorrect and require remediation
2. Validate these assumptions against archived DVLA file information if it is available

Stage 3 builds on Stage 2, but will be ran for the cases where the system assumptions could not be applied (e.g. if the audit trail was incomplete). Stage 3 will validate these cases against the DVLA archives to see how many records show the symptoms of the data error, and therefore require remediation.

6. Key Assumptions for Remediation of Stage 3 Records

The only system assumptions that applies to records for Stage 3 is as follows:

1. The system assumptions cannot determine whether a record is correct or incorrect, i.e. from a system assumption perspective, the record is labelled as “data position unknown”. This is usually because older records are lacking a full audit trail, or a recent DVLA update feed was received which could have been either partial or full.
2. Only records where the last update is based on a DVLA(GB) source file should be considered

Stage 3 also relies heavily on the following business assumption for such records:

3. When ODR records are found in the source DVLA(GB) archives, and a comparison between the DVLA archive record and the ODR record indicates that the data transposition error has occurred, then the record will be selected for remediation.

These assumptions will be applied to the remaining 604,171 records coming out of Stage 2.

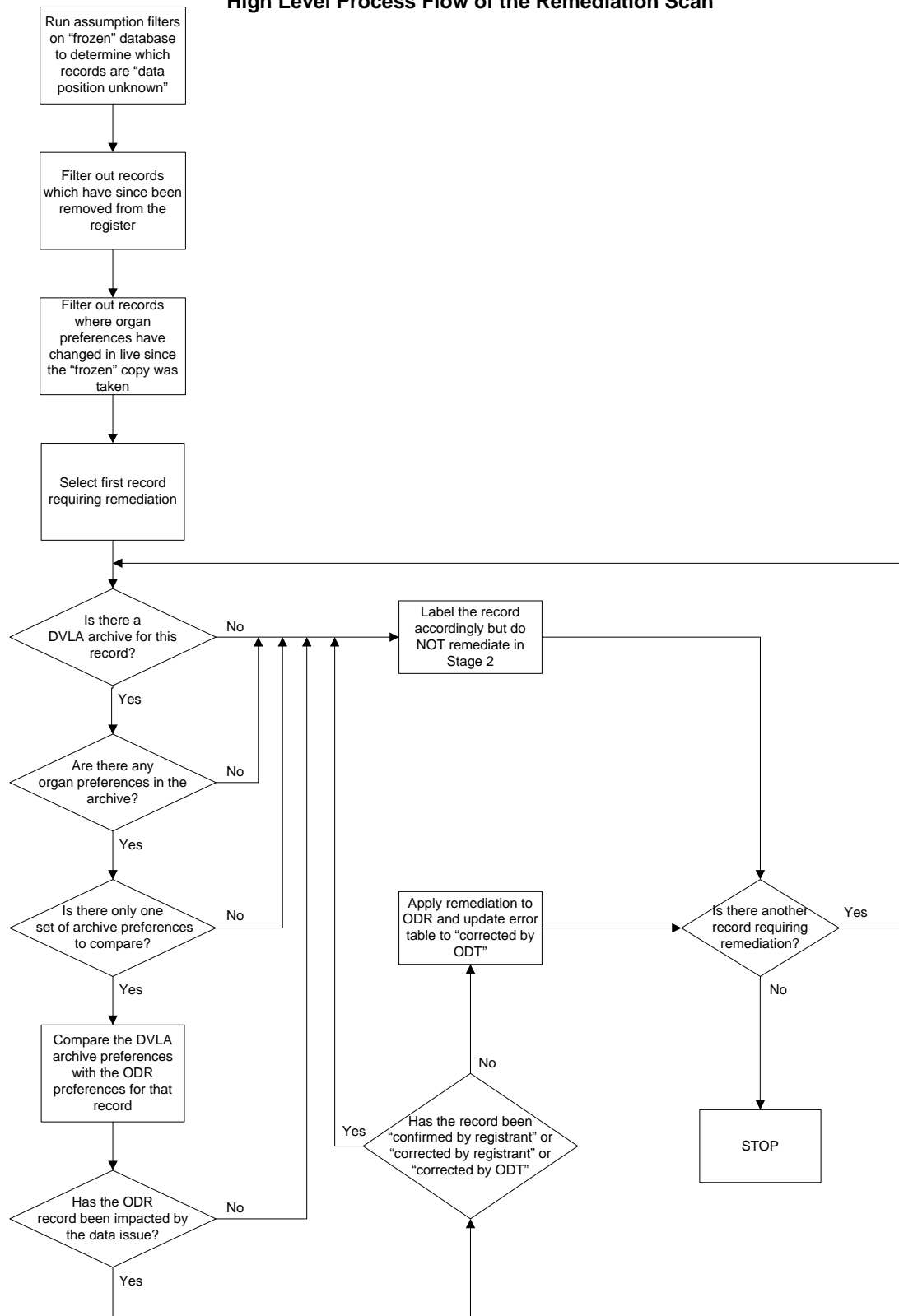
7. Validation of Assumptions

Of the remaining 604,171 records from Stage 2, 276,040 records were identified as “data position unknown” based upon the system assumptions. Once these were compared with the DVLA archive files, 101,708 displayed the characteristics of the data error.

The approach for Stage 2 therefore is to remediate records where the system assumptions cannot work out the data position, but where matching with the DVLA archives indicates that the data error is occurring. This means remediating 101,708 records. Other records will be left unchanged, and continue to be flagged as “unknown”, and passed through to Stage 4 for consideration.

By signing off this document, all approvers are agreeing to the remediation of the 101,708 records based on the logic defined in this document. The process that will be followed for this remediation is covered off in other documentation, but a high level process flow is attached below for reference.

High Level Process Flow of the Remediation Scan



ODR Organ Preferences DVLA Data Issue

***Approach & Assumptions
to the Analysis of Impacted Records on the ODR
Stage 4
DE1699-N4***

3 May 2010

V1.2

Table of Contents

1. Summary of changes	3
2. Reviews and Approvals.....	3
3. Introduction.....	4
4. Purpose of Document.....	4
5. Data Analysis Approach for Stage 4	4
6. Key Assumptions for Remediation of Stage 4 Records	5
7. Validation of Assumptions.....	5

1. Summary of changes

This section records the history of significant changes to this document. Only the most significant changes are described here.

Version	Date	Author	Description of change
0.1	21/04/10		Draft for discussion with 3 approvers
1.0	23/04/10		Version for sign off
1.1	27/04/10		Updated Version for sign off (section 7 revised)
1.2	03/05/10		Updated Version for sign off (better explanation of rationale & risks added)

Where significant changes are made to this document, the version number will be incremented by 1.0.

Where changes are made for clarity and reading ease only and no change is made to the meaning or intention of this document, the version number will be increased by 0.1.

2. Reviews and Approvals

Approved by:

Name	Role	Signature	Date
David Shute	Assistant Director – Transplant Support Services		
Alan McDermott	Director of Business Transformation Services		
Ian Bateman	Assistant Director of Quality		

3. Introduction

An investigation of the organ preferences problem by the IT department resulted in a Problem Statement being documented, which records the coding error that has led to incorrect organ preferences records being loaded onto the ODR. In summary, the data problem is:

- For DVLA(GB) records received since 1999, where specific organ preferences have been requested, the preferences for the following organs have been transposed:
 - a. The registrants Liver preference has been recorded as their Lungs preference
 - b. The registrants Lungs preference has been recorded as their Cornea preference
 - c. The registrants Heart preference has been recorded as their Liver preference
 - d. The registrants Cornea preference has been recorded as their Heart preference

The data analysis is being conducted in a series of stages – this document follows on from Stage 1 of the Data Analysis, which ring-fenced 997,014 records as being potentially impacted by the error, and Stages 2&3, which remediated a total of 494,551 records which were validated against DVLA files as definitely incorrect, and identified 130,530 records as likely to be incorrect but which cannot be validated. This left a total of 371,933 records for further analysis.

4. Purpose of Document

This document summarises the approach to Stage 4 of the analysis of the ODR records that have been impacted by the fault in processing DVLA(GB)'s organ preference feed. A key part of this approach involves making assumptions about the data on the ODR, and validating those assumptions when necessary.

The approach needs to be agreed by 3 key people:

1. David Shute on behalf of the ODR business
2. Alan McDermott as IT Director
3. Ian Bateman in his role of Quality Assurance

This document has been drafted to document the approach and facilitate agreement and sign-off of the approach.

5. Data Analysis Approach for Stage 4

The analysis approach that is proposed for Stage 4 is to determine ODR records that can be proven as NOT impacted by the data problem. Achieving this was done in 2 main steps:

1. Determine a set of system assumptions that can be used to derive a subset of the ring fenced records that are correct and do not need to be ring fenced
2. Validate these assumptions against archived source file information if it is available

6. Key Assumptions for Remediation of Stage 4 Records

The system assumptions that were applied to records for Stage 4 are as follows:

1. Feeds from FHSA, Scottish FHSA, Online Registrations, DVLA(NI), MPL & NICSA feeds always send through organ preferences that would overwrite, and therefore correct, any prior data received from the DVLA(GB).

Evidence for assumption: Analysis of archive files from the above source feeds has shown that these feeds always send organ preferences. Over 5m archive records were checked to confirm this assumption. Testing of other feeds has confirmed that the transposition problem only occurs on the DVLA(GB) feed.

2. Feeds from Boots, Citizen Card, Goldfish & NSTS Tracing Service feeds never send through organ preferences, so the preferences remain unchanged when these feeds are processed.

Evidence for assumption: Analysis of archive files & code from the above source feeds has shown that these feeds never send organ preferences.

3. The DVLA(GB) transposition problem did not occur before 01.01.99, so any record not updated since 1999 can be considered as correct

Evidence for assumption: Analysis of the database shows that a step change in the percentage of registrants agreeing to donate eyes occurred in early 1999. There is no business reason for expecting a step change at this time. In addition, there was an upgrade to the database platform implemented in early 1999 which required a rewrite of the DVLA processing scripts. Finally, a sample (c.50) of records originating from 2000 onwards was checked with the DVLA after the problem was discovered, and all displayed the transposition problem.

These assumptions were applied to the remaining records for analysis, and 197,601 of these records were deduced to be correct. This confirms that the ring fence set up at Stage 1 was a 'rough and ready' ring fence (based upon 3 key assumptions only) and that further refinement of the system assumptions was able to demonstrate more records as being correct.

7. Validation of Assumptions

Of the remaining records coming out of Stage 3, 197,601 records were identified as actually correct based upon the system assumptions. These were then validated against the archived source file information where it was available. Unlike Stages 2 and 3, the validation of these records was more complex, as an exact match against archive files and ODR was not required. Instead, the purpose of this validation was to prove that the Stage 4 records were not subject to the DVLA transposition error, so a two step validation method was applied as below:

1. Validation was first conducted against the non-DVLA(GB) archive source files where they were available. This step checks to see if the ODR preferences were correct as per the expected source of the last update. Records could fail this validation for a variety of reasons, and not necessarily because of the DVLA transposition problem, so these records were passed onto a second validation.
2. For records which failed the above validation, the second validation check was to compare the ODR preferences against the last DVLA(GB) archive, and see if the preferences resembled the DVLA(GB) transposition problem as it is described in the Problem Statement.

This two step approach was required as the purpose of Stage 4 was not to prove that the ODR preferences matched the non-DVLA(GB) archives, but that the ODR preferences are not impacted by the DVLA(GB) transposition problem, as set out in the Problem Statement. If the first step validation only was conducted, the error rate would include all types of data quality problems, system & processing errors, and not just the transposition problem that is the subject of this remediation.

Of the 197,601 records assumed to be correct, 98,588 were found in the non-DVLA(GB) archives, and 97,361 were confirmed as exact matches between the archive files and the current ODR preferences, and can therefore be confirmed as correct. Note that records were matched against the appropriate archive depending on their latest source type. For example, if the latest source feed was from the FHSA, these records were validated against the FHSA archive files alone.

An additional 1,227 records did not show an exact match between the ODR record and the non-DVLA(GB) archive. However, this does not necessarily mean that the record has incurred the DVLA(GB) data problem. There are a number of different possible explanations to a non-match, such as unaudited (but correct) updates being applied to the ODR, missing source archive files, records being skipped and not processed, or even other (as yet unidentified) system problems. To establish whether this was the case, the 1,227 records were then ran through a comparison with the DVLA(GB) archive, to see if any displayed the symptoms of the DVLA(GB) problem. 587 of the 1,227 records were matched against a DVLA(GB) record, and only 11 displayed the DVLA(GB) error as defined in the Problem Statement.

It should be noted that although the 11 show the patterns of the transposition problem, they may not have been caused by the DVLA coding error. The system records show that the last update on these records was not from the DVLA, so the 11 that showed the patterns of the DVLA problem could be down to registrants changing their preferences in a particular way that mimics the DVLA problem or different ways of collecting and recording information by the different feeds, as well as the random chance of the ODR flags hitting the expected error pattern. However, erring on the side of caution, the worse case was considered where all 11 were assumed to be caused by the DVLA transposition error, and an error rate therefore calculated based on these results extrapolated across the whole 197,601 population where system assumptions conclude the records to be correct.

When extrapolated, the 11 errors out of 587 in the second validation step equates to an estimated 23 errors out of the 1,227 records passed to the second validation step. When combined with the results from the first validation step, this equates to an estimated error rate of 0.023%, or 46 records out of the total 197,601 sample size.

As this is an estimated error rate, the likely range around this figure was requested from the Stats team assuming (i) that all 1,227 records incurred the same error rate as the 587 records matched, and (ii) a more cautious estimate assuming the remainder of the 1,227 records incurred twice the error rate of the 587 records matched. With a similar error rate, the 95% interval estimate for the number of errors in the whole 197,601 sample is from 27 to 65 (0.014% to 0.033%). If the error rate doubles for the unmatched records, the 95% interval becomes 47 to 93 (0.024% to 0.047%).

A business decision was therefore required as to whether to remove only records that have been positively validated out of the ring fence, and therefore leave the remaining c.100,000 records as 'unknown', or to accept the low error rate, and remove all 197,601 records from the ring fence. Principally this involved choosing between the following 2 risks:

- Removing around 100,000 registrants from the ODR register, and therefore not being able to follow the wishes of c.100,000 registrants who have previously registered organ donation preferences with NHSBT. Although these registrants will be subject to the mailing exercise which requests confirmation of their preferences, feedback from the initial mailings has indicated that less than 10% are responding to the letters. Where no response is received, the ODR preferences are not regarded as a trusted source of the registrants' wishes, and not passed on to next of kin.
- Accepting the risk that an estimated error rate of 46 records on the ODR could be incorrect due to the transposition problem. It is highly unlikely these would convert into likely donors (46 records equates to 0.14 donors whose wishes would not be accurately recorded on the ODR - based on 575 donors out of 179,817 ODR removals. Using the more conservative estimates and upper limit of the 95% confidence range, 93 incorrect records translates to 0.3 donors). This is before making any allowance that over time, the records could be updated by new feeds and therefore corrected anyway.

A decision was made with the 3 approvers on Thursday 29 April that the accepting the second risk was preferable. By signing off this document, all approvers are agreeing to the confirmation that 197,601 no longer need to be flagged as "unknown" records based on the logic defined in this document.

SUI WEB TRAFFIC & ODR REGISTRATIONS ANALYSIS 10 April-1 May					
Date	Total visits	Unique visitors	Total Page Views	ODR Web Registrations	ODR Web Withdrawals
Sun 4 Apr	1,064	1,011	3,953	580	7
Mon 5 Apr	1,412	1,312	5,261	776	7
Tue 6 Apr	2,766	2,631	11,329	1,504	22
Wed 7 Apr	3,510	3,317	13,866	1,845	20
Thu 8 Apr	2,398	2,252	9,381	1,843	18
Fri 9 Apr	1,909	1,785	7,385	1,042	15
Sat 10 Apr	1,577	1,504	6,099	595	7
Weekly total:	14,636	13,812	57,274	8,185	96
Sun 11 Apr	9,685	9,340	40,077	6,003	177
Mon 12 Apr	3,871	3,665	15,106	2,404	52
Tue 13 Apr	3,572	3,360	13,954	1,800	34
Wed 14 Apr	3,124	2,887	12,195	1,842	21
Thu 15 Apr	2,382	2,203	8,861	1,263	15
Fri 16 Apr	1,942	1,797	7,433	1,017	8
Sat 17 Apr	1,240	1,150	5,000	658	8
Weekly total:	25,816	24,402	102,626	14,987	315
Sun 18 Apr	1,424	1,305	5,479	701	11
Mon 19 Apr	2,159	1,985	7,816	994	14
Tue 20 Apr	5,167	4,903	19,368	2,948	26
Wed 21 Apr	2,381	2,264	8,703	1,388	9
Thu 22 Apr	2,028	1,889	7,441	1,021	15
Fri 23 Apr	1,756	1,649	6,357	867	14
Sat 24 Apr	1,168	1,097	4,288	660	7
Weekly total:	16,083	15,092	59,452	8,579	96
Sun 25 Apr	1,460	1,393	5,207	819	8
Mon 26 Apr	2,146	2,002	7,796	1,115	13
Tue 27 Apr	2,029	1,913	7,612	1,127	10
Wed 28 Apr	1,804	1,702	6,550	969	14
Thu 29 Apr	1,605	1,510	5,783	847	17
Fri 30 Apr	1,317	1,218	4,721	771	3
Sat 1 May	877	824	3,229	533	8
Weekly total:	11,238	10,562	40,898	6,181	73
Total:	67,773	63,868	260,250	37,932	580
SUMMARY					
6,003 registrations on Sun 11 Apr vs 580 on Sun 4 Apr. Increase of 5,423 (+1035%)					
2,404 registrations on Mon 12 Apr vs 776 on Mon 5 Apr. Increase of 1,628 (+310%)					
9,685 total visits on Sun 11 Apr vs 1,424 on Sun 18 Apr. Decrease of 8,261 (-680%)					
6003 registrations on Sun 11 Apr vs 701 on Sun 18 Apr. Decrease 5,302 (-856%)					
Registrations remained high until Fri 16 Apr. Withdrawals decreased after Wed 14 Apr. Spike in increases on 20 Apr - possibly due to TV programme					
Withdrawals decreased by 219 (70%) & 242 (77%) in weeks 2&3 after SUI.					

ORGANDONATION.NHS.UK WEB TRAFFIC & ODR REGISTRATIONS ANALYSIS 10-18 Apr					
Date	Total visits	Unique visitors	Total Page Views	ODR Web Registrations	ODR Withdrawals
Sat 10 Apr	1,577	1,504	6,099	595	7
Sun 11 Apr	9,685	9,340	40,077	6,003	177
Mon 12 Apr	3,871	3,665	15,106	2,404	52
Tues 13 Apr	3,572	3,360	13,954	1,800	34
Wed 14 Apr	3,124	2,887	12,195	1,842	21
Thur 15 Apr	2,382	2,203	8,861	1,263	15
Fri 16 Apr	1,942	1,797	7,433	1,017	8
Sat 17 Apr	1,240	1,150	5,000	658	8
Sun 18 Apr	1,424	1,305	5,479	701	11
Totals:	28,817	27,211	114,204	16,283	333
SUMMARY					
6,003 registrations on Sun 11 Apr vs 580 on Sun 4 Apr. Increase of 5,423 (+1035%)					
2,404 registrations on Mon 12 Apr vs 776 on Mon 5 Apr. Increase of 1,628 (+310%)					
9,685 total visits on Sun 11 Apr vs 1,424 on Sun 18 Apr. Decrease of 8,261 (-680%)					
6003 registrations on Sun 11 Apr vs 701 on Sun 18 Apr. Decrease 5,302 (-856%)					
Mon 12 Apr - Thur 15 Apr:					
Hits, registrations and withdrawals continued to rise until Thur with a decrease from Fri 16 Apr.					

ORGANDONATION.NHS.UK WEB TRAFFIC & ODR REGISTRATIONS ANALYSIS 27 Mar-14 Apr									
Date	Total visits	Unique visitors	Average time on site (seconds)	Total Page Views	ODR Web Registrations	ODR Withdrawals	ODR Web Withdrawals		
Sat 10 Apr	1,577	1,504	3.39	6,099	595	7	Sun 4 Apr		7
Sun 11 Apr	9,685	9,340	3.48	40,077	6,003	177	Mon 5 Apr		7
Mon 12 Apr	3,871	3,665	3.58	15,106	2,404	52	Tue 6 Apr		22
Total					9,002		Wed 7 Apr		20
(Easter weekend)							Thu 8 Apr		18
Sat 3 Apr	1,137	1,065	3.52	4,196	632		Fri 9 Apr		15
Sun 4 Apr	1,064	1,011	3.39	3,953	580	7	Sat 10 Apr		7
Mon 5 Apr	1,412	1,312	3.50	5,261	776	7	Weekly Total:		96
Total					1,988				
Sat 27 Mar	2,952	2,831	3.12	9,925	1,594		Sun 11 Apr		177
Sun 28 Mar	2,589	2,448	3.48	9,753	1,390		Mon 12 Apr		52
Mon 29 Mar	3,638	3,420	3.28	12,811	1,758		Tues 13 Apr		18
Total					4,742		Wed 14 Apr		43
							Week to date Total:		290
SUMMARY									
Sun 11 Apr:									
9,685 total visits compared to 1,064 & 2,589 on previous 2 Sundays									
9,340 new unique visitors to the website									
6,003 total registrations from the website compared to 580 & 1,390 on previous 2 Sundays									
40,077 total page views									
Mon 12 Apr:									
3,871 total visits compared to 1,412 & 3,638 on previous 2 Mondays									
3,665 new unique visitors									
2,404 total web registrations compared to 776 & 1,758 on previous 2 Mondays									
15,106 total page views									
Withdrawals									
177 withdrawals on Sun 11 Apr & 52 on Mon 12 Apr.									
We have checked through the reasons given and of the 290 withdrawals for Sun 11 to Wed 14 Apr, 28 can be attributed directly to the media coverage of the data issue and 87 to change of mind. The remainder are a mix of reasons that cannot be attributed to the incident.									

Identification of change in DVLA registration pattern

By comparing the pattern of registrations on the ODR amongst those who were not registered through the DVLA with the pattern of DVLA registrants, any change in pattern in DVLA registrations over the years can be identified. This analysis is restricted to those who express a specific wish for donating a heart, lungs, liver or corneas. Those who express a wish to donate 'all organs', or who expressed a wish to donate kidneys or a pancreas, are not included in the analysis, although they are included in the total number of patients in Table 2.

The first step is to consider all those who have not initially registered through DVLA. We can then obtain the percentage of this group who express a willingness to donate each organ for each year since 1994 (first year of DVLA registrations). These figures are given in Table 1 below.

Table 1: Percentage of non DVLA registrants willing to donate heart, lungs, liver and corneas.

Year	Heart	Lungs	Livers	Corneas
1994	75	88	91	41
1995	75	83	90	36
1996	81	85	91	49
1997	75	81	89	32
1998	75	79	88	35
1999	76	88	92	33
2000	75	82	89	42
2001	73	82	89	37
2002	72	81	89	36
2003	74	80	88	36
2004	74	82	89	34
2005	75	83	89	32
2006	76	82	89	30
2007	77	82	90	29
2008	78	83	90	26
2009	80	84	92	22
2010	85	90	95	12
Overall	76	83	90	35

While there is some variation in the percentages willing to donate each particular organ across the years, there is no evidence of a systematic change. The overall percentages, for each organ, will therefore be used in subsequent analysis.

Now suppose that these overall percentages (last line of table) apply to all DVLA registrants in each year. We can then calculate the expected number of registrations for each organ, based on the percentage non-DVLA registrants, for each year. If there is a sudden change in the pattern of DVLA registrations, compared to what is expected, this would indicate that a change in coding may have occurred.

Table 2 gives the numbers of DVLA registrations in each year, the observed numbers for each organ, and the expected numbers based on non DVLA registrations.

Table 2: Observed and expected numbers of DVLA registrants willing to donate heart, lungs, liver and corneas.

Year	Hearts		Lungs		Liver		Corneas		Total
	Obs	Exp	Obs	Exp	Obs	Exp	Obs	Exp	
1994	5434	6029	6289	6584	6880	7140	3454	2777	7933
1995	43046	50734	53260	55407	57723	60080	26252	23364	66755
1996	37263	44265	46304	48342	49917	52419	22288	20385	58243
1997	23247	28930	28902	31595	31243	34259	14747	13323	38066
1998	23124	29454	29514	32167	31573	34880	15525	13564	38755
1999	261	365	439	398	436	432	281	168	480
2000	55367	86214	90683	94155	91296	102096	60215	39704	113440
2001	20282	38531	41707	42080	40494	45629	29451	17745	50699
2002	20654	40348	43389	44064	41835	47780	30568	18581	53089
2003	18011	37777	40697	41256	38568	44735	29251	17397	49706
2004	24618	57036	61712	62289	58586	67542	46207	26266	75047
2005	25034	60957	65100	66571	61634	72185	49533	28072	80206
2006	24393	63249	67676	69074	63537	74900	53117	29128	83222
2007	22244	57758	62741	63078	58285	68397	49542	26599	75997
2008	19232	53582	59387	58517	53768	63453	50067	24676	70503
2009	20690	57556	64596	62858	58223	68159	56168	26506	75732
2010	5195	14271	16243	15586	14769	16900	13814	6572	18778

The number of registrations processed in 1999 is much smaller than previous years but this is balanced by the large number in 2000.

Some changes in the actual DVLA registration patterns, compared to what would have been expected from other registrations, can also be seen. In particular, the observed number of heart registrants becomes very much lower than expected after 1999, while the observed number of cornea registrations becomes greater than what would have been expected. This suggests that a coding issue began around this time. This pattern is in agreement with what would have been expected if cornea registrations were coded as something else while heart was coded as cornea.

A measure of 'discrepancy' between observed and expected values is obtained for each year by summarising the values $(\text{observed} - \text{expected})^2 / (\text{expected})$ across all organs in each year. By dividing by (expected) we take account of the fact that differences based on larger expected numbers of observations will have greater variation. We then sum this measure, across the four organs, for each year to give an overall measure of discrepancy between DVLA registrations and what would have been expected from non-DVLA registrations.

The values of this measure of discrepancy are shown in Table 3.

Table 3: Values of discrepancy measure that summarises the extent of differences between the observed and expected values in each row of Table 2.

Year	Discrepancy
1994	247
1995	1697
1996	1490
1997	1764
1998	2176
1999	110
2000	22904
2001	16947
2002	18095
2003	19277
2004	34757
2005	39151
2006	45380
2007	43123
2008	49639
2009	58303
2010	14048

The main message from this table is that the value of the discrepancy measure suddenly increases after 1999. This provides clear statistical evidence of a change in registration pattern after 1 January 1999.

There is some discrepancy between observed and expected DVLA registrations before 1999. This may well be due to different demographic make up of DVLA registrants; in particular there will tend to be rather more younger people in this group.

March 2010