Evaluation of data-matching pilots: confirmation

Interim findings

December 2012

Executive summary

In June 2011 the Government announced plans for a fundamental change to the way we register to vote, moving from a system of household to individual electoral registration (IER). The change is designed to address vulnerabilities in the current system and enable people to take responsibility for their own vote.

The current Government proposals include the intention to compare existing electors' names and addresses on the electoral register with records held by the Department for Work and Pensions (DWP) in order to verify the identity of these people on the register. This matching process is known as 'confirmation'.

Electoral Registration Officers (EROs) from 12 local authorities in England and Wales and two Scottish Valuation Joint Boards piloted the confirmation process from July – December 2012. The Electoral Commission has a statutory responsibility to report on the effectiveness of these pilots. Our full report will be published in March 2013 - this paper presents the emerging findings in order to inform the Government's policy choices around the implementation of IER.

The initial results from comparing electoral register entries with the DWP database show:

- A significant number of entries on the selected electoral registers[1], **70.6%** in total, can be matched with the records on the DWP-Customer Information System database.[2] This is a figure for the pilot areas and does not represent a national match rate.
- There are significant variations across pilot areas: the match rate varies from **55.2%** in Tower Hamlets to **82.6%** in Wigan. The variations are likely to be due to a range of factors including differences in demographics between the areas and our full report will include further analysis on this.
- There is no evidence from these pilots that areas with a lower proportion of green matches have significantly less accurate electoral registers - as we do not know which database (the register or the DWP database) is up to date in any given case.
- There are significant variations between different types of electors. The match rate for **postal voters** was **78.6%** and **80.7% for**

---

[1] Those in force as of July 2012.
[2] By matched, we mean those register entries which received a green rating from the matching process. See Section 2 of our main report for details of the green, amber, red ratings.

attainers.[3] However, the match rate was very low for some groups, such as students.

- When the exercise was repeated following the annual canvass, the match rate increased to **75.1%.**[4] This figure is likely to reduce once the final results are available due to the addition of London boroughs, with their more mobile population.

Initial findings from interviews with participating local authorities and a review of the pilot process show that:

- While the headline findings are positive there are a number of additional issues which should be considered alongside the pilot results.
- These pilots did not test a process or system that would be used in live running, in particular the IT system. We are therefore unable to comment on whether there is yet a solution for undertaking confirmation that is suitable for use by all EROs across England, Scotland and Wales.
- Because of this, we are also unable to estimate resource or training implications for local authorities. However, EROs with lower match scores are likely to require more resources than others because any electors who cannot be matched will need to be written to and canvassed in order to collect their personal identifiers. The variation in the level of match scores achieved in these pilots is therefore significant.
- Where an ERO received a non-perfect (amber) match, they were limited in the action they could take in part because they were not able to view the names and addresses held by DWP.
- There was significant variation between the register formats of local authorities, meaning that reformatting work had to be undertaken before the registers were processed by DWP. Without this re-formatting, it is possible that the match rate would have been lower.[5]

Conclusions, recommendations and next steps:

- Overall, the interim results suggest that the 'confirmation' process provides a reliable way of verifying the identity of people on the electoral register and we **recommend** that the policy should be pursued as part of the transition to IER and further development work should be carried out on the processes and resource implications for all EROS.

---

[3] An attainer is someone who will turn 18 before the end of the twelve month period starting from the next 1 December after their application for registration is made.
[4] This figure applies to only those areas which had Police and Crime Commissioner elections and therefore concluded their canvass early.
[5] Further testing is planned in order to explore this issue in December and January.

- There is an urgent need for testing of the process for undertaking confirmation, in particular the transfer of data between local authorities and DWP, the time required and any particular skill or resource needs. Otherwise there is a potential for inconsistency in the implementation of confirmation across England, Scotland and Wales.
- The higher level of match score post-canvass shows this is the most accurate match and we **recommend** that during the transition to IER, confirmation should take place as soon as possible following the conclusion of a canvass.
- Further work is also required to get greater certainty on the expected overall green match score. For example, it is possible that the match score will reduce once data which has not been reformatted is matched against the DWP database. In addition, the figures from these pilots do not represent a national match rate for Great Britain and further analysis is required to explore how the national rate might differ from these results.
- We **recommend** investigation of options for increasing the proportion of electoral register entries that can be matched, as this will make the transition to IER simpler for voters and more cost-effective.

Our full evaluation, to be published in March 2013, will also look at:

- The results from the post-canvass matching for those pilot areas that did not have elections for Police and Crime Commissioners in November 2012 and did not therefore run an earlier canvass.
- Costs and resources of the pilot scheme for all organisations involved.
- Evidence collected from electoral administrators with regards to the potential impact of data-matching on their work and resources.
- An analysis of the results from using locally-held data for matching.
- Any further testing undertaken by the Cabinet Office Electoral Registration Transformation Programme team.

1 Background/overview

1.1 The current Government proposals for the introduction of individual electoral registration (IER) include the intention to compare existing electors' names and addresses on the electoral register with records held by the Department for Work and Pensions (DWP) in order to verify the identity of people currently on the register. This process is known as 'confirmation'. All those people whose entries are not confirmed, or those who wish to make a new application or to vote by post or proxy, will be asked to provide unique identifying information – name, National Insurance number and date of birth.

1.2 This new aspect of the policy is based on findings from the first round of data matching pilots in 2011, run by the Cabinet Office and evaluated by the Electoral Commission.[6] These pilots tested the usefulness of providing electoral registration officers with access to data held by other public bodies in order to improve the accuracy and completeness of their registers. An interesting, but untested findings from the 2011 pilots suggested that up to 68% of entries on the register could be verified through a match with the DWP database. Our evaluation report therefore stated that there was a need for more testing and evidence on the confirmation proposal to check its reliability.

1.3 This second round of pilots was designed to test the effectiveness of confirmation and particularly the accuracy of the matching process used. There are 14 areas involved in the pilots across Great Britain and the piloting is taking place between July and December 2012.

1.4 This interim evaluation, and data in this paper, is based on partial results from those authorities taking part in the pilot.[7] Data is still being analysed and may be subject to changes. Our full statutory evaluation, incorporating the results from all pilot authorities, will be available in March 2013.

2 Process

2.1 The pilots were overseen by the Cabinet Office Electoral Registration Transformation Programme team, working closely with the Department for Work and Pensions (DWP), who managed the process for comparing entries on the participating authorities' electoral registers with data in the Department

---

[6] The Electoral Commission, *Data matching schemes to improve accuracy and completeness of the electoral registers – evaluation report* (March 2012) http://www.electoralcommission.org.uk/__data/assets/pdf_file/0010/146836/Data-matching-pilot-evaluation.pdf

[7] Pre-canvass results are available for all areas but post-canvass results are only available for those areas that held an earlier canvass as a result of holding Police and Crime Commissioner elections in November 2012.

for Work and Pension - Customer Information System database (DWP-CIS). Government Digital Service (GDS) were responsible for building the IER IT system that carries the data from local authorities to DWP. The Electoral Commission was responsible for providing a statutory evaluation of the pilots.

2.2   The matching process is carried out using an algorithm created by DWP that compares entries in the two datasets:

1. The process starts by comparing the Unique Postcode Reference Number (UPRN) and, where not available, the different address lines.
2. Once two properties have been matched, the algorithm compares the data relating to the individuals – first name, last name, middle name (initials or full name depending on the register's format) and date of birth (only for attainers) – held on the two databases.[8]
3. The algorithm creates a 'best match' field between the different data for the individual.
4. GDS transform the 'best match' into a colour (RAG status):
    a. **Green**: the register entry has been matched positively[9] with the DWP-CIS database;
    b. **Amber**: the register entry has been matched but not entirely, the individual cannot be confirmed with certainty;
    c. **Red**: the register entry has not been matched.
5. A summary file, providing totals for the number of register entries in each RAG category and a full match file indicating the RAG category for each individual on the register were then sent back to the pilot authorities.
6. Some pilot authorities carried out some further checks using locally-held data to verify the accuracy of the matching. However, some authorities did not analyse the data either due to a lack of know-how or to the timing with Police and Crime Commissioner (PCC) elections taking place during the same period.

2.3   The Commission was involved in the development and approval of the matching algorithm. The matching process employed for this set of pilots was significantly more sophisticated than that used for the first wave of pilots, with the involvement of staff from DWP with a high level of expertise in this work. A small number of the pilot areas also spent considerable time helping to refine the matching algorithm.

---

[8] An attainer is someone who will turn 18 before the end of the twelve month period starting from the next 1 December after their application for registration is made.
[9] A positive match is where two entries have been fully matched or with a "fuzzy" match: for example two entries with a Full Last Name, Full Middle Name and Fuzzy First Name match have been marked as green. Fuzzy matches include examples where there is a small difference in spelling or where one name sounds like another or where one name includes another, e.g. 'Thea' and 'Althea'. The vast majority of matches were generated by a "Full First Name – Full Last Name" match (58.9%) and "Full First Name – Full Last Name – Middle Name Initial" match (33.9%).

3 Initial results

3.1   The matching exercise was carried out twice: once using the registers in force before the annual canvass, and a second time with the new registers published after the canvass. Due to the Police and Crime Commissioner (PCC) elections taking place across England and Wales (excluding London)[10] the matching against the post-canvass register was conducted in two phases.

3.2   It is important to note that the 14 areas, despite being geographically spread, do not constitute a representative sample of local authorities or population throughout England, Scotland and Wales. As such the data presented here should not be read as applicable to Great Britain as a whole.

3.3   All data presented below are provisional at this stage.

Table 1: Pre and Post-canvass matching approach.

Area	Pre-canvass	Post-canvass
PCC election (England excluding London, Wales)	July electoral register matched against July DWP-CIS database. Results available for all areas.	**October** electoral register matched against **October** DWP-CIS database. Results available now.
No PCC election (London, Scotland)		**December** electoral register matched against **December** DWP-CIS database. Results available in mid-December.

Key findings

3.4   The results show:

1.   **Match rate.** A total green match rate of **70.6% for the pre-canvass** register and **75.1% for the post-canvass** register.[11]
2.   **Significant variations from local authority to local authority.** The match rates vary considerably across areas. In the pre-canvass matching, green scores ranged from 55.2% in Tower Hamlets to 82.6%

---

[10] Local authorities with PCC elections had to anticipate their annual canvass. The results for the post-canvass matching for non-PCC areas will be available in mid-December.
[11] The post-canvass results are for the areas with PCC elections only.

of Wigan. These variations are likely to be due to a range of demographic and other factors. For example, London boroughs and other cities have key factors that may be affecting the accuracy of the register and/or the DWP-CIS database, and therefore their match rates. These include significant population mobility and a younger population, including students. Our full report will include further analysis on the impact of the differences between areas on their match rates.

3. **Pre-canvass and post canvass.** Except for Manchester and Ceredigion, the post-canvass match rates are higher than the pre-canvass. This was anticipated as the post-canvass register is more accurate and complete than the pre-canvass register[12]. It is unclear why Manchester and Ceredigion do not follow the pattern set by the other pilots and further analysis is required. However, data for the post-canvass matching are not available for all areas and the total (75.1%) is expected to go down with the inclusion of the results from the London boroughs.

4. **First round of pilots (autumn 2011) against second round (autumn 2012).** The match score is slightly higher than the one observed in the first wave of data-matching in 2011 (68%). However there are too many factors to be taken into consideration to assess, at this stage, if the higher match rate is due to an improvement with the algorithm and the process, to the demographics of the piloted areas or to differences in the methodology adopted. We will assess this further in our full evaluation.

5. **Amber matches.** The matching returned a low rate of amber matches: a total of 3.1% for the pre-canvass and 2.9% for the post-canvass. Considering the low number of records in this category, it may be worth reviewing the scoring and value of this category. Our full report will contain further analysis on these amber matches and the implications for EROs.

6. **Red matches.** A total of 26.3% of entries in the pre-canvass register and 22% of those in the post-canvass were classified as red. Of these, just over 4% could not be matched because the addresses for the entries could not be matched.

7. **Particular groups**. An initial review of the evidence suggests that the process of confirmation is not effective for some groups, including students and some others in large communal accommodation such as care homes. In the case of students this seems to be because data held by DWP tends to relate to their home address rather than their term time address. Further work will be required to reach these groups.

---

[12] We estimate that the completeness of the registers declines, on average, by around one percentage point a month from the completion of the annual canvass (The Electoral Commission, *The completeness and accuracy of the electoral registers in Great Britain*, March 2010).

Green matches

3.5   Table 2 displays the percentage of entries on the pilot area electoral registers that received a green rating when compared against entries on the DWP-CIS.

Table 2: Pre-canvass and post-canvass green matches.

Local Authority	Pre-canvass green matches	Post-canvass green matches[13]
Ceredigion	60.7%	59.9%
Conwy	76.5%	78.4%
Greenwich	71.4%	-
Harrow	74.3%	-
Lothian	68.9%	-
Manchester	61.3%	61.3%
Peterborough	78.6%	79.4%
Powys	75.9%	77.8%
Renfrewshire	74.0%	-
Southwark	58.8%	-
Sunderland	79.5%	82.5%
Tower Hamlets	55.2%	-
Wigan	82.6%	84.7%
Wolverhampton	78.2%	80.4%
Total	**70.6%**	**75.1%**
Average	**71.8%**	**75.6%**

Amber matches

3.6   A register entry is marked with the amber RAG status when a name held by the register and the DWP database at an address cannot be clearly demonstrated to relate to the same person. This is either because there is only a 'fuzzy' match with either first name or surname or where the register entry matches against more than one entry in the DWP database – 'multiple matches'.

3.7   Table 3 shows the percentage of amber matches for each participating authorities.

Table 3: Pre-canvass and post-canvass amber matches.

---

[13]   For those areas holding PCC elections only.

Local Authority	Pre-canvass amber matches	Post-canvass amber matches[14]
Ceredigion	2.3%	2.3%
Conwy	2.6%	2.6%
Greenwich	2.6%	-
Harrow	3.5%	-
Lothian	2.8%	-
Manchester	3.0%	3.0%
Peterborough	2.9%	2.9%
Powys	1.8%	1.8%
Renfrewshire	4.0%	-
Southwark	4.2%	-
Sunderland	3.6%	3.8%
Tower Hamlets	4.5%	-
Wigan	2.4%	2.5%
Wolverhampton	3.0%	3.2%
Total	**3.1%**	**2.9%**
Average	**3.1%**	**2.7%**

3.8   The percentage of all register entries classified as amber is 3.1% for the pre-canvass register and 2.9% for the post-canvass register. There is small variation across authorities, ranging from 1.8% to 4.5% in the pre-canvass registers.

3.9   It is worth noting the impact of 'multiple matches' on the amber classification. Our analysis found that the large majority of entries marked as amber, 60.4% in the pre-canvass register, were classified in this way due to multiple matches with the DWP database. In many of these cases the match itself was strong (e.g. an exact first name, last name match). We will investigate this issue further in our full evaluation.

Red matches

3.10 Table 4 shows the percentage of red matches in the pre-canvass and post-canvass register for each local authority.

3.11 The percentage of red matches for the pre-canvass register is 26.3% and 22% for the post-canvass one. As with the green matches, there are significant variations across participating authorities with again Wigan and Tower Hamlets at opposite ends of the scale – 15% and 40.4% respectively.

---

¹⁴ For those areas holding PCC elections only.

3.12 It should be noted that there is no evidence from these pilots that areas with a larger proportion of red matches have significantly less accurate registers - as we do not know which database (the register or the DWP) is up to date in any given case.

Table 4: Pre-canvass and post-canvass Red matches.

Local Authority	Pre-canvass red matches	Post-canvass red matches[15]
Ceredigion	37.1%	37.8%
Conwy	20.9%	19.0%
Greenwich	25.9%	-
Harrow	22.3%	-
Lothian	28.3%	-
Manchester	35.6%	35.7%
Peterborough	18.6%	17.8%
Powys	22.4%	20.4%
Renfrewshire	22.1%	-
Southwark	36.9%	-
Sunderland	16.9%	13.7%
Tower Hamlets	40.4%	-
Wigan	15.0%	12.9%
Wolverhampton	18.8%	16.4%
Total	**26.3%**	**22.0%**
Average	**25.8%**	**21.7%**

3.13 Just over 4% of red matches are a result of a failed address match Excluding the register entries at the properties that failed to match, the percentage of red matches would be 21% in the pre-canvass (against 26.3% of all register entries). There is therefore some scope for improving the level of matching by improving the address matching rules. For further details on the process of address matching see the section below.

Address match

3.14 The electoral register is an address-based database as an individual's register entry is connected to a specific property. A fundamental part of the matching algorithm is devoted to finding an address match between the register and the DWP-CIS; without it, entries cannot go forwards for further matching and will automatically receive a red rating.

---

[15] For those areas holding PCC elections only.

3.15 The address was matched in two ways, firstly by matching a Unique Property Reference Number (UPRN), if this was available. Many, although not all, electoral registers make use of UPRNs but it has not historically been included on the DWP-CIS.

3.16 Our evaluation of the first round of data-matching recommended the inclusion of UPRNs on the DWP database but at the start of the pilots, the DWP database did not have UPRNs assigned to its records. UPRNs were added to some of the DWP data used for the pilots but due to the time available, not all properties in the pilot areas had UPRNs assigned and the levels varied between areas. In the absence of a UPRN, the matching process compares the individual address lines on the register and DWP-CIS.

3.17 The results for the pre-canvass registers show that a total of 76.8% register records had their address matched via UPRN, 18.8% via the address matching process and 4.4% of records were unable to have their addresses matched. There is some variation in these results between the pilot areas which merits further investigation.

3.18 It is likely that if the address match rate can be improved the overall green match score will increase. For example, there is some evidence that the use of the word 'Flat' in an address makes it less likely to match. Further analysis is required but it should be possible to develop the matching rules in order to take account of this anomaly.

3.19 Maximising the level of address matching is critical as there is some evidence, from initial analysis carried out by DWP, that many register entries at unmatched addresses could be successfully matched against a DWP record if the addresses can be matched first.

Postal voters, proxy voters and attainers

3.20 Certain entries on the register, postal voters, proxy voters, and attainers, are flagged as such on the register. Below we present the green match results broken down by these categories.

3.21 The analysis is limited as data are not available for all local authorities[16].

Postal and proxy voters

3.22 Under the current proposals for implementing IER, postal voters who are not confirmed and do not supply their personal identifiers will lose their absent vote when the new register is published on 1 December 2014.

3.23 The results show:

---

[16] Data for Lothian are missing for all flags.

1. **Green match rate.** The overall match rate for postal voters in the **pre-canvass register was 78.6%** (compared with 70.6% for all entries)**.**
2. The **match rate for postal voters is higher than** the one achieved with **all registers' entries**. Further analysis is required but we expect that this is at least partly due to the demographic profile of postal voters, who tend to be older and move less frequently than the general population.

3.24 The overall match rate for proxy voters was **64.4% in the pre-canvass register**. As with postal voters, further analysis is required to understand why this difference occurs.

Attainers

3.25 Our analysis found an overall match rate for attainers of **80.7% in the pre-canvass register**. This is encouraging as there have been some concerns about the risk of low registration amongst attainers under IER.

3.26 The matching process for attainers also included the elector's date of birth. Due to problems in the matching process, the date of birth was only matched correctly for three local authorities: the date of birth match rate between the electoral registers and DWP-CIS was 88.4%.[17]

4 Verifying the results

4.1   In order to determine how accurate the matching process has been, we have taken two approaches.

Pre/Post canvass match file comparison

4.2   We compared each entry in the pre-canvass match results against the post-canvass register (where available) in order to determine if the elector on the pre-canvass register was still resident after the canvass. By doing so, and on the assumption that the register published right after the canvass is at its best for completeness and accuracy, it is possible to assess the quality of the pre-canvass matching.[18]

4.3   The results for authorities which held PCC elections (and could therefore provide a post-canvass register at this point) showed that **94.3% of entries**

---

[17] We are currently investigating the reasons behind the problems with matching electors' date of birth.
[18] Previous research by the Commission has shown that levels of accuracy and completeness are highest immediately following the conclusion of the canvass. We estimate that the completeness of the registers declines, on average, by around one percentage point a month from the completion of the annual canvass (The Electoral Commission, The completeness and accuracy of the electoral registers in Great Britain, March 2010).

classified as green before the canvass were also green after the post-canvass matching.[19] This suggests that the green matching results can be regarded as trustworthy and there were no significant issues with the matching algorithm incorrectly matching records.

Dip-sampling

4.4   We asked those pilots without PCC elections (who were therefore not completing their canvass until later) to randomly select a sample of the pre-canvass register entries matching each category (green/amber/red) and to check whether the same names could be found in responses to the annual canvass. The size of the sample varies across authorities.

4.5   The results of the dip-sampling exercise are still being collated at the time of writing and further analysis of the data supplied by the local authorities is required. However, provisionally, the data indicates that the large majority of individuals on the pre-canvass register that were marked as green were found during the canvass. Also, it appears that a significant proportion of individuals on the pre-canvass register marked in the red category were found by EROs on their registers after the canvass. Our full report will look at these results in more detail.

5   Costs and resources

5.1   At this stage, we are unable to assess the costs of the pilot schemes (see below for further information). However, we can make the following comments:

- Some elements of the system used in the pilots may be useable if the process is rolled out nationally. The costs of these elements of the system will therefore not relate to the pilots only. We will explore with the Cabinet Office how the costs can be broken down to allow for an assessment of whether the scheme resulted in saving of time and costs as per our statutory evaluation.
- For the same reasons, the resources required for this work within the pilot areas do not provide a reliable benchmark for what would be required for full implementation.
- EROs in areas – likely the ones with a young and mobile population – will have to carry out more work (for example, issuing letters, using door-to-door canvassers) than those with higher match rates to follow up those electors who cannot be confirmed.

---

[19] This figure excludes those records which were carried forward from the pre-canvass register to the post-canvass register (i.e. where the person did not actively respond to the canvass), where that information was made available. This is done in order to reach as strong an assessment as possible of the accuracy of the matching. If all records on the post-canvass register that were carried forward are included in this analysis the 94.3% increases to 97%

6   Emerging issues

6.1   At the time of writing, we are still to conclude interviews with all participating authorities and the other organisations involved. In addition, the matching process is still underway for areas without PCC elections and so data may be subject to review. Nevertheless, there are a number of emerging issues for consideration.

Local authority confidence in the matching algorithm

6.2   Although the electoral registers were matched against DWP-CIS data, the pilots did not receive any DWP-CIS data back with their match results. The match file simply assigned a RAG match status to each entry on their electoral register. This was due to concerns from DWP around the supply of personal data.

6.3   Practically, this means that the pilots had to accept the RAG status assigned to a record without having the ability to see clearly why it has been assigned. This is arguably less of an issue for green matches where the 'Best match' code will indicate the nature of the match (see Appendix A), for example an exact match on first name and surname.

6.4   However, electoral administrators from several pilot areas have raised concerns that, without having sight of the DWP-CIS data, they do not have the confidence to confirm any entries which did not receive a green rating.

6.5   For both amber and red matches, the absence of the detail from the DWP records meant that using locally-held data was the only method by which pilots could make an assessment of the matching beyond the RAG status offered. For example, many records that matched amber will have done so because more than one good match exists on the DWP database. Without access to the names held by DWP, EROs would need to consult other local data in order to make a judgement about whether one of those matches is good enough to justify retaining a register entry.

6.6   In our full evaluation, we will consider further how EROs should interpret the data they receive, based on feedback from participating pilots. It will be important to ensure that a consistent process to how people are added to and removed from the register.

Electoral Registers' format

6.7   When electoral registration data was collected to send to DWP, a significant variation between pilot authorities' register formats was discovered.

In order to test as robustly as possible how well the matching algorithm was working, the Cabinet Office took a decision to ask Government Digital Services (GDS) to manipulate the data into a useable format before matching.[20] At this stage it is not known what impact this may have had on the data but we would expect that without this manipulation, the match scores would have been lower. It is also unclear what cost will be associated with any manipulation that is required.

6.8   The Cabinet Office will repeat the matching process using unmanipulated data in January, and we will assess the impact of this activity in our full evaluation. However, it is possible either that local authorities will need to undertake data cleansing or standardisation activity in advance of any confirmation test, or that there will need to be data manipulation undertaken by a third party when rolled out. These options will have both staff capacity and cost implications.

Data transfer between local authorities and DWP

6.9   The data transfer processes used in the confirmation pilots were not the method that would be used in normal operation. In live running, the intention is that data will move automatically and securely from the local authority into the IER system and then onto DWP, where the register data will be matched and then sent back to authorities. We understand that the testing of this process will take place between January and April 2013.

6.10 For the pilots, the IER system had not yet been secured and the links had not been built between the local authority system and the IER system, or the DWP system and the IER system. Therefore, an interim solution of transferring electoral registers via secure courier was employed. While secure, this was time-consuming and costly for all those involved and is unlikely to be a sustainable solution for long term running.

6.11 The 'live-running' data transfer processes were not tested during the pilots due to two factors. DWP had not yet developed the link between their system and the IER system within the proposed timetable; this work is now scheduled for January.

6.12 During the pilots, it was also discovered that the systems that hold electoral registers (Electoral Management Software systems) are not currently linked into the main government secure network (GCSx). This connection is required because the IT approach, to delivering both confirmation and IER in general, assumes that these EMS systems are able to directly interface with the central IER system. The development of this connection for a small number of the pilot authorities is planned for December and January, but a wider exercise needs to be undertaken to assess the scale of the issue. It is

---

[20] For example, to ensure that address and name fields were correctly separated into the fields that the IER system expected to receive.

unlikely that a single solution will be appropriate for all local authorities and this will have implications for how the system is rolled out.

6.13 This issue means that there has not yet been a full test of the process and we are unable to comment on whether the proposed live running process will be suitable for use by all EROs across the country. We are also unable to ascertain the resource implications for local authorities at this stage. The project should expect that further issues may arise in relation to data transfer and sufficient time to resolve these issues needs to be built into the timetable for delivering the system.

Cost and resources

6.14 At this stage, we do not have any reliable information on cost or resource implications. However, given the variation in green matches, it looks likely that certain EROs will require more resources than others to undertake the follow up of electors required after the confirmation exercise. It is important that this requirement is built into funding allocations.

Roles of the organisations involved

6.15 Currently, the data-matching system is heavily reliant on third parties (DWP and GDS) managed through the ERTP programme within Cabinet Office. Some issues arose, during the period of these pilots, around the lack of clarity of roles and delivery to deadline, particularly around the delivery of the IT system. The process needs to be ironed out through clear working practices and service level agreements and progress is being made in this area. The working practices also need to clearly outline who owns the data as it travels through each part of the system.

7   Conclusions and next steps

7.1   As this is an interim evaluation, all conclusions at this stage are subject to validation from the full pilot results when available in the New Year.

Confirmation match results

7.2   The pilots have revealed to date that the confirmation process provides a reliable way of verifying the identity of people on the electoral register and we recommend that the policy should be pursued as part of the transition to IER and that further development work should be carried out on the processes and resource implications for all EROS.

7.3   The higher level of match score post-canvass shows this is the most accurate match and we recommend that during the transition to IER,

confirmation should take place as soon as possible following the conclusion of a canvass.

7.4   Further work is required on the matching algorithm, in particular taking on board feedback from pilot authorities to resolve any issues which may be specific to particular electoral register formats and to refine the green, amber and red match categories.

7.5   Further work is also required to get greater certainty on the expected overall green match score. For example, as outlined above, it is possible that the match score will reduce once unmanipulated data is matched against the DWP database. In addition, the figures from these pilots do not represent a national match rate for Great Britain and further analysis is required to explore how the national rate might differ from these results.

7.6   Options for increasing the overall green match rate should also be investigated, such as what steps an ERO may need to take to amend the format of their data and the full integration of UPRNs into both DWP and electoral registers.

Confirmation process

7.7   Far less can be said at this stage about the process local authorities will follow to verify electors' identity via the confirmation process. These pilots, contrary to the recommendation of our previous evaluation, did not test the process or system that would be used in live running. This means that we are unable to comment at this stage on the costs, resource or skill implications for local authorities.

7.8   The data transfer process in particular requires further testing. This is essential not only for confirmation, but also when the full IER system is in place, as new applications and those entries not verified will need to pass through this system.

7.9   We recommend that once ready, testing with the system planned for roll-out should be carried out in advance of distribution to local authorities. This testing should include:

- Testing the data transfer mechanism between local authorities, the main IER system and DWP

- Assessing the time taken to carry out the confirmation process

- Providing results in a standardised format for feedback from local authorities

- identifying any particular additional skills or resource needs for local authorities

7.10 We understand that the Cabinet Office plan to undertake this testing from January to April 2013, subject to delivery of key parts of the system from

suppliers. The window for development is reasonably small but this testing should enable an assessment on the deliverability of the new IER system according to current plans. We will report on any additional information as part of our full evaluation.

Appendix A

Table 1A: Match rate by RAG status and pilot authority.

Local Authority	Pre-canvass			Post-canvass		
	Green	**Amber**	**Red**	**Green**	**Amber**	**Red**
Ceredigion	60.7%	2.3%	37.1%	59.9%	2.3%	37.8%
Conwy	76.5%	2.6%	20.9%	78.4%	2.6%	19.0%
Greenwich	71.4%	2.6%	25.9%	-	-	-
Harrow	74.3%	3.5%	22.3%	-	-	-
Lothian	68.9%	2.8%	28.3%	-	-	-
Manchester	61.3%	3.0%	35.6%	61.3%	3.0%	35.7%
Peterborough	78.6%	2.9%	18.6%	79.4%	2.9%	17.8%
Powys	75.9%	1.8%	22.4%	77.8%	1.8%	20.4%
Renfrewshire	74.0%	4.0%	22.1%	-	-	-
Southwark	58.8%	4.2%	36.9%	-	-	-
Sunderland	79.5%	3.6%	16.9%	82.5%	3.8%	13.7%
Tower Hamlets	55.2%	4.5%	40.4%	-	-	-
Wigan	82.6%	2.4%	15.0%	84.7%	2.5%	12.9%
Wolverhampton	78.2%	3.0%	18.8%	80.4%	3.2%	16.4%
Total	**70.6%**	**3.1%**	**26.3%**	**75.1%**	**2.9%**	**22.0%**
Average	**71.1%**	**3.1%**	**25.8%**	**75.6%**	**2.7%**	**21.7%**

Table 2A: Best match field.

Match combination	Identity RAG
Date of birth (DOB), LAST NAME, FIRST NAME, MIDDLE_NAME	Green
DOB, LASTNAME, FIRSTNAME, MIDDLE_NAME INITIAL	Green
DOB, LASTNAME, FIRSTNAME FIRST 3 INITIALS, MIDDLE_NAME	Green
DOB, LASTNAME, FIRSTNAME	Green
DOB, LASTNAME, FIRSTNAME FIRST 3 INITIALS	Green
DOB, LASTNAME, FIRSTNAME FIRST 3 INITIALS, MIDDLE_NAME INITIAL	Green
DOB, LASTNAME, FUZZY FIRSTNAME, MIDDLE_NAME	Green
DOB, LASTNAME, FUZZY FIRSTNAME, MIDDLE_NAME INITIAL	Green
DOB, LASTNAME, FUZZY FIRSTNAME	Green
DOB, LASTNAME, MIDDLE_NAME	Green
DOB, LASTNAME, MIDDLE_NAME INITIAL	Green
DOB, FUZZY LASTNAME, FIRSTNAME, MIDDLE_NAME	Green
DOB, FUZZY LASTNAME, FIRSTNAME, MIDDLE_NAME INITIAL	Green
DOB, FUZZY LASTNAME, FIRSTNAME	Green
DOB, FUZZY LASTNAME, FIRSTNAME FIRST 3 INITIALS	Amber
DOB, FUZZY LASTNAME, MIDDLE_NAME	Amber
DOB, FUZZY LASTNAME, MIDDLE_NAME INITIAL	Amber
DOB, FUZZY LASTNAME	Amber
DOB, LASTNAME	Green
LASTNAME, FIRSTNAME, MIDDLE_NAME	Green
LASTNAME, FIRSTNAME, MIDDLE_NAME INITIAL	Green
LASTNAME, FIRSTNAME	Green
LASTNAME, FIRSTNAME FIRST 3 INITIALS	Green
LASTNAME, FUZZY FIRSTNAME, MIDDLE_NAME	Green
LASTNAME, FUZZY FIRSTNAME, MIDDLE_NAME INITIAL	Green
LASTNAME, FUZZY FIRSTNAME	Amber
LASTNAME, FIRSTNAME FIRST 3 INITIALS, MIDDLE_NAME	Amber
LASTNAME, FIRSTNAME FIRST 3 INITIALS, MIDDLE_NAME INITIAL	Amber
FUZZY LASTNAME, FIRSTNAME, MIDDLE_NAME	Green
FUZZY LASTNAME, FIRSTNAME, MIDDLE_NAME INITIAL	Green
FUZZY LASTNAME, FIRSTNAME	Amber
FUZZY LASTNAME, FIRSTNAME FIRST 3 INITIALS	Amber
FIRSTNAME AND LASTNAME REVERSED	Green