Working Paper

Predicting the outcome of Disability Living Allowance (DLA) adult claims

by Diana Kasparova, Karen Mackinnon
and David Wilkinson

DWP Department for Work and Pensions

Department for Work and Pensions

Working Paper No 99

Predicting the outcome of Disability Living Allowance (DLA) adult claims

Diana Kasparova, Karen Mackinnon and David Wilkinson

A report of research carried out by the Policy Studies Institute and the National Institute of Economic and Social Research on behalf of the Department for Work and Pensions

This document/publication is also available on our website at:
http://research.dwp.gov.uk/asd/asd5/rrs-index.asp

Any enquiries regarding this document/publication should be sent to us at:
Department for Work and Pensions, Commercial Support and Knowledge Management Team,
Upper Ground Floor, Steel City House, West Street, Sheffield S1 2GQ

First published 2011.

ISBN 978 1 84712 939 0

Contents

List of tables

Acknowledgements

Throughout the project we received a huge amount of help and support from numerous officials at the Department for Work and Pensions (DWP). We especially want to thank the project managers Mike Daly, Michelle Harrison and Andrea Kirkpatrick who provided excellent help and guidance to ensure the project was completed successfully.

We also thank Alan Marsh and Jean Martin for their expert advice throughout the project.

We also benefited from feedback and advice from an Advisory Group convened by the DWP. The project relied on the provision of a large number of claim packs and on checking the ability of coded claim packs to convey the information contained in the original packs. We therefore are extremely grateful to the Disability and Carers Service (DCS) team at DWP and their decision makers for their support and contribution that was crucial to the success of the project.

We also thank the Brentwood coding office at National Centre for Social Research for developing the coding system for claim packs and we are indebted to Group Sigma Ltd for their diligent and speedy coding of all claim packs at short notice.

The Authors

Dr Diana Kasparova was a Senior Research Fellow at the Policy Studies Institute at the time of the project.

Karen Mackinnon is a Research Fellow at the Policy Studies Institute specialising in data analysis and large-scale data management. She has considerable experience of handling complex survey and administrative data.

David Wilkinson is a Senior Research Fellow at the National Institute of Economic and Social Research. His work focuses largely on evaluation of welfare-to-work programmes as well as the impact of recent changes in student funding.

Abbreviations

CCM	Customer Case Management
DCS	Disability and Carers Service
DLA	Disability Living Allowance
DM	Decision Maker
DWP	Department for Work and Pensions
EMP	Examining Medical Practitioner
GPFR	General Practitioner Factual Report
HFR	Hospital Factual Report
PDCS	Pensions, Disability and Carers Service

1 Background and introduction

The purpose of this report is to see how well coded data from Disability Living Allowance (DLA) adult claim packs can predict the outcome of these claims. The work forms part of a project designed to assess the feasibility of developing a survey instrument that would be able to collect information equivalent to that contained in the DLA claim pack and present it in a more structured form than free-text format. The project is part of an overall programme of work exploring the possibility of estimating the take-up of DLA, see Kasparova *et al.* (2007) for more details.

Previous work has coded data from 622 DLA adult claim packs[1] and we have looked at how a simple classification of customers based on a limited number of coded questions from the main claim form can predict claim outcomes. The work presented here briefly recaps that existing work and extends it to cover the use of coded data from additional evidence forms.

In Chapter 2 we briefly review the coding exercise and recap on earlier work to look at the potential usefulness of some of the coded data from additional evidence forms. Chapter 3 then goes on to explore how this work can be used to predict the probability of care and mobility awards. The first part of this section reviews how we derived variables that attempt to replicate DLA care and mobility eligibility criteria, then we present some descriptive analysis based on these derived variables and finally report models for the probability of receiving a care or mobility award. Chapter 4 presents our conclusions.

[1] The current version of the DLA application form is very similar to the one used in this research. It can be found at http://www.dwp.gov.uk/advisers/claimforms/dla1a_adult.pdf

2 The coding exercise

A selection of claim packs was made available from the Disability and Carers Service (DCS)[2] based on achieving a distribution of packs across all types and combinations of award (higher, middle, lower and nil rate Care and higher, lower and nil rate Mobility). The target was to have 240 claim packs, but for operational reasons we received just 200. We used 120 claim packs to develop a coding frame for the information included in the claim packs and then using these coding frames were able to code the remaining 80 claim packs.

2.1 Forms used for coding

Table 2.1 shows the forms included in these 200 claim packs, indicating that each pack contained four documents on average. Early analysis of some of these packs identified the 'core' documents as being a claim form (Customer Case Management, CCM, based and dated April 2006 or April 2007), DBD39 (information gathering) and DBD43 (Decision and data processing record). Later on, information on the award rates and evidence used in decision making were more likely to appear in the DBD810 and the ENT1 forms. Of these documents only the claim form contained evidence on care and mobility needs presented by the customer; the remaining forms contained information submitted by decision makers explaining the reasons behind their decisions. The most frequent type of evidence additional to that contained in the claim pack was the General Practitioner Factual Report (GPFR) form and the documents that DMs sought least frequently related to Special Rules (DS1500 and DBD520) and mental health (DBD365M) cases.

Coding pro-forma were designed for each type of document identified above, containing a detailed description of the fields that needed coding. The selection of information into each pro-forma followed a careful examination of the contents of each document in terms of their relevance to the claiming and decision making processes.

Some documents, such as prescriptions and care plans, had to be left out of the coding because of the complex and highly individual nature of the data.

Open-ended questions allowed customers to answer in free-text format, i.e. they gave the customers full freedom to say what they wanted and in the words they chose to do so. Coding frames for open-ended questions had to be designed by a specialist coding organisation. Success in developing code frames for open-ended questions depended on the availability of a sufficient number of documents where a field to be coded contained information. However, while a claim form was present in each pack, other documents were not so common and even within a document some questions were rarely answered.

Where a total number of cases containing information that had to be coded (say, 'Why no evidence' question in DBD385) was smaller than 50, a code frame for the open-ended question could not be constructed. However, since some code frames were transferable among similar questions across a number of documents certain fields were coded even in the documents that were not present in large numbers.

[2] The DCS has recently been renamed as the Pensions, Disability and Carers Service (PDCS).

Table 2.1 Summary of documents coded

Document	Document description	Number	%
CCM (2006/2007)		200	100
CCM2006	Claim Form used in 2006	66	33
CCM2007	Claim Form used in 2007	134	67
DBD810	Decision form recording DLA award	116	58
GPFR	General Practitioner Factual Report	90	45
DBD43	Decision and data processing record	80	40
DBD39	Registration and information gathering	76	38
DBD508/DBD33	Request for additional information from claimant – report of telephone call, request for call back or interview report	43	22
DBD385	Medical Services advice request/evidence sheet	33	17
ENT1	Notification (letter) to customer of DLA decision and details of any award made	31	16
IB85	Incapacity for Work – Medical Report Form	20	10
DBD365G (HFR)	Hospital Factual Report (c.f. GPFR)	18	9
DBD520	Special Rules medical advice to DM	14	7
EMP	DLA Examination Report – Doctor's assessment with reasons for referral to doctor from DM	10	5
DBD378N	Community Psychiatric Nurse/Mental Health Report	8	4
DS1500	Doctor's report for DLA to accompany a patient's claim under Special Rules	6	3
DBD365M	Factual Report – Mental Health (c.f. HFR)	5	3
DBD520MS	DLA/AA Special Rules Advice - response from Medical Services	3	2
Total packs		200	100
Total documents		753	

The coding pro-forma was shaped by supporting documentation received from DCS for two documents – the ENT1 and the IB85 form. The ENT1 form is a letter sent to the customer outlining the outcome of their claim and the reasons for the decision. Its pro-forma was constructed from a template used by DMs and called 'Notification Output'. Since this template contained all possible variations of the letter a customer might receive, this enabled the research team to construct a coding pro-forma and all code frames for this form.

The IB85 form is a medical report form filled in by doctors for customers claiming Incapacity Benefit. If a customer has undergone such a report it is present in the customer's claim pack as a 'supporting evidence' document. It contains large amounts of complex and detailed medical information, which is tailored to a particular medical issue that the report aims to clarify. When designing the coding pro-forma containing the possible values for each question in IB85, the researchers drew on Annex 2 (Descriptors and Scores in Each Functional Area) and Annex 3 (Mental Health descriptors) to '*The medical assessment of Incapacity Benefit*' (Benefits Agency 1994). These descriptors appeared to be used methodically within the IB85 form (electronic version) and this allowed a pro-forma to be constructed and a range of categories to be designed for the variables.

2.2 Verification of coded data by Decision Makers

The coding discussed above was applied to 622 claim packs and 494 of these packs were verified by DMs. Our initial intention was for the same DM to verify the coded claims as made the original decision, so as to minimise any 'noise' from idiosyncratic variation in DM behaviour. Unfortunately, however, for operational reasons this was not possible. DMs were asked whether they recalled the case based on the coded claim from only, and asked again whether they recalled the case having seen the coded additional evidence. Just four DMs recalled the case based only on the coded claim form. However, two of these then decided that they did not recall the case once they had seen the coded additional evidence. A further five DMs recalled the case only once they had seen the coded additional evidence.

The verification exercise consisted of requests to DMs to arrive at a decision only using information from the coded packs and then they were asked to consider additional evidence contained in the coded claim pack and arrive at a decision again.

At the beginning DMs were given coded claim forms and asked to tick the pieces of information they judged to be crucial in making an award decision. They were also asked a brief set of questions based on the coded claim forms:

1 Do you recall the case?

2 Are you able to make a decision on the basis of information up to this point?

3 If able to make a decision, what will be the rate for both Care and Mobility components of DLA based only on the coded information contained on the claim form.

Their awards at this point constituted decision 1.

The DMs were then asked to view the coded additional evidence, see Table 2.2, (if there was any) and asked to tick the coded information for each piece of evidence if they thought it was crucial in making a judgement on the award. After assessing all the evidence DMs were asked the same set of questions outlined above.

Their awards at this point constituted decision 2.

For cases where there was only a claim form, DM responses at decision 1 were carried forward to decision 2.

The following forms were included in the additional evidence available to DMs at decision 2:

- IB85.
- DBD33.
- DBD365G (HFR).
- DBD365M.
- DBD385.
- DBD520.
- EMP.
- GPFR.

We were then able to create variables indicating whether the same decision was made for the original claim and the coded claims.

It is a basic feature of the approach used that we would not, in the course of our work, be able to say which decisions were 'right' and which were 'wrong'. And it was an important message to the DMs taking part in the verification exercise that we were not seeking to check the reliability of their work, but the accuracy of the coding system which we had developed. Nevertheless, for ease of exposition, in the following discussion we refer to a decision made by a DM as 'right' if it matched the original decision, 'wrong' otherwise'.

Then by comparing the 'real' outcome with the 'verified' outcome at both decision points, we created four new variables:

- If at decision 1 the DM recorded on the verification data that they thought the customer should be awarded Higher rate Care and on the real data the customer was indeed awarded DLA Care at the Higher rate then this was coded as a 'correct' decision at Decision point 1.
- If the DM then went onto change their mind at Decision 2 and recorded that the case only warranted Middle rate Care, then this would be coded as an 'incorrect' decision at Decision point 2[3].

Similar variables for the Mobility component were calculated comparing the decisions at each of the two decision points with the 'real' decision made originally on the claim pack.

2.3 Usefulness of additional evidence

We then consider the usefulness of additional evidence in making 'correct' decisions. It is of course possible that cases with more pieces of additional evidence are more difficult to make decisions about, so some caution is needed when interpreting the usefulness of the additional evidence.

Table 2.2 shows the percentage of verified claims that DMs felt able to make a decision on based on the coded claim form data. Here, DMs only felt able to decide on just 84 (17 per cent) claims, with a slightly higher number being able to make a judgement on mobility than care. The majority of these decisions agreed with the real world decision (70 per cent of care decisions and 80 per cent of mobility decisions).

Table 2.2 Decision 1 – made after seeing coded claim form data only

		Of which able to decide, per cent				
	Total packs	Total	On care	On mobility	Of which correct decisions, %	
					On care	On mobility
Total packs	494	84	81	84	57	67
Out of which:		17%	16%	17%	70%	80%

Table 2.3 looks at the decisions when the coded additional evidence was available for consideration. Here DMs felt able to decide on 227 (46 per cent) claims. Here again we find that the majority of decisions agreed with the real world decision (69 per cent of care decisions and 81 per cent of mobility decisions).

[3] We did not include any coded version of any prescriptions contained in the claim pack.

Table 2.3 Decision 2 – made after seeing coded claim form data and additional evidence

Documents in the pack	Total packs	Of which able to decide, per cent				
		Total	On care	On mobility	Of which correct decisions, per cent	
					On care	On mobility
Total Packs	494	227	213	215	147	173
Out of which:		46%	43%	44%	69%	81%
Claim form but no additional evidence	59	35	35	35	30	32
		59%	59%	59%	86%	91%
Claim form and one additional evidence document	194	86	79	83	52	63
		44%	41%	43%	66%	76%
Two additional evidence documents	149	72	67	69	42	58
		48%	45%	46%	63%	84%
Three additional evidence documents	68	27	27	23	19	16
		40%	40%	34%	70%	70%
Four or more additional evidence documents	24	7	5	5	4	4
		33%	20%	20%	80%	80%

The rest of Table 2.3 shows the number and percentage of 'correct' decisions by the amount of additional evidence available. Some general observations can be made on the basis of these tables.

First, the fewer rates available on the mobility component compared with the care component are likely to make the adjudication easier and lead to fewer wrong decisions. Furthermore, within the mobility component of DLA, the higher and lower rates are for fundamentally different types of condition. The higher rate is for physical difficulties in walking, and the lower rate is for need for supervision. Therefore, there is very little possibility of awarding the lower rate when it should be the higher rate or vice versa.

Second, the very high proportion of correct decisions made on cases where in reality there was only a claim form to support the claim (second row) suggests that in these cases the customers have made a clear case on their claim form alone. More importantly, it implies that our coding conveyed all the necessary information for DMs to judge on the case correctly in 86 per cent and 91 per cent of care and mobility awards respectively. These proportions are likely to represent a maximum for the goodness of fit of our models predicting the probability of award.

Third, for the 57 and 67 care and mobility decisions on which verifying DMs could decide correctly at the decision 1 point (see Table 2.2), in real life only 30 and 32 had all the necessary information on the cases presented on the claim form alone (see Table 2.3, second row). In the remaining 27 and 35 instances, verifying DMs were able to make correct decisions without using the evidence that had been requested and used in real life. This reflects the fact that we asked the verifying DMs to make a judgement if they possibly could, even if they would, for preference, have asked for more evidence. Some of this additional evidence was of a nature that did not add information but was simply required for procedural purposes[4]. In some cases, where higher rate mobility awards are considered, DMs are obliged to seek additional information before making a decision, even if they think it unnecessary.

[4] Only one of the 35 cases had only a DBD385 form, which is a request for further information.

However, the majority of evidence forms did contain additional information, so the ability of DM's to make a correct decision without the additional evidence available suggests that in some cases it is possible to arrive at a correct probability of award on the basis of the limited (coded) information presented on the claim form alone.

Fourth, the fact that the same proportions of right decisions were achieved in both cases (after seeing the claim form alone 70 per cent care decisions and 80 per cent mobility decisions were correct and after seeing additional evidence the respective figures were 69 per cent and 81 per cent) may suggest that we are seeing the level of natural noise in the system that is due to the judgement present in the decision making process, probably in addition to the noise created by the loss of information caused by coding. Thus, our multivariate models may not be expected to be able to correctly predict claim awards for more than this percentage of cases.

The noise here comes from the same DMs arriving at different decisions on the same real claim pack and/or from different DMs deciding on the real and coded case[5] and/or from the reaction of DMs to coded information as opposed to real information. The main contribution of additional evidence is that it enables DMs to arrive at a decision – the proportion of those who were able to decide at Decision 2 is much higher than the proportion of those who decided at Decision 1.

Excluding the claims with no additional evidence, a breakdown of the overall figures by the number of pieces of additional evidence (see Table 2.3, last four rows) shows that packs with up to two pieces of evidence were the easiest to decide. The proportions of DMs able to decide on care and mobility rates were 45 per cent and 46 per cent respectively. As cases were becoming more complex and requiring more additional information the proportions of DMs able to decide on the case were lower. These cases may just be too complex to make a decision about without extremely detailed information. In relation to our multivariate modelling approach, it may be extremely difficult to get the required detailed information for these complex cases and our models may not be good at predicting claims where lots of additional evidence is required.

In the coded forms the information that claim packs were able to convey was less complete than in the real form and the degree of this incompleteness grew as cases became more complex. However, where DMs could make a decision when they had more evidence they were better able to make the same decision based on the coded data as the real world decision for care awards (63 per cent of decisions were the same based on two pieces of additional evidence compared with 70 and 80 per cent for three and four pieces of additional evidence). The same was not true for mobility awards, where the percentage with the same prediction were broadly similar irrespective of the number of pieces of additional evidence.

2.4 Types of evidence

By analysing the extent to which decisions based on coded data, after taking into account additional evidence, were different from the real world decision we can identify what pieces of evidence were most useful in making a decision. The evidence that is most likely to be associated with wrong care and wrong mobility rates include:

- DBD33.
- GPFR.
- DBD365G.

[5] We do not know the proportion of DMs that decided on someone else's case, but know that few verifying DMs recalled cases, as mentioned in Section 2.2.

- DBD378(N).
- IB85.
- Examining Medical Practitioner (EMP).
- DBD365(M).
- DBD385.

Given that some of the forms were not coded (e.g. DBD33) and some were not fully coded (e.g. GPFR)[6], some of the wrong decisions can be attributed to the reduced information available to DMs as a result of the coding and/or poor presentation of the information available on the original evidence (e.g. some GPs submitted poor or illegible information). Among the types of evidence listed above, the only ones that seemed to be associated with the ability of DMs to make a decision (but not necessarily a right decision) included the IB85 and the DBD378(N), suggesting that their coding did not significantly impair the volume of information that they contained in real life. In these cases the reasons for wrong decisions may be attributed to cases being genuinely difficult. For example, the presence of a DBD385 in the list above indicates that in real life a DM has requested further information on the case. However, since this list of evidence covers almost all types of evidence, except DBD520 and DS1500, definitive conclusions on the usefulness of particular types of additional evidence are difficult to make.

The role of the DBD520 and DS1500 forms is revealed by the analysis of cases where DMs changed their decisions between Decision 1 and Decision 2 from wrong to correct. In these cases both of these forms were likely to be present. The majority of DBD520s were related to malignant disease and this explains the presence of DS1500 forms in packs containing a DBD520. The DS1500 is completed by a health professional when the customer is claiming under Special Rules as a result of terminal disease. The type of information recorded on this form suggests that the simple presence of this form might have signalled to DMs that a terminal illness had been diagnosed or queried.

This analysis suggests that our research should focus on the content of the evidence rather than its type when evaluating the usefulness of evidence. This is because the same documents may contain different information, e.g. in one case a GPFR answers DMs' questions and in another case not. Another reason is that different documents may contain the same information, e.g. information on needs is asked across a number of documents (such as GPFR and EMP).

2.4.1 Health conditions

Customers record their illness and disabilities on the claim form and we used the DWP coding frame of conditions/disabilities in order to code these conditions. We then created a 'grouped' or 'collapsed' version of the illnesses/disabilities suffered by customers as described in Table 2.4. The grouping fit well with the impairment classification described by Purdam *et al.* (2008) giving support to the developed classification[7]. Each variable describing the grouped health condition (from disabl1r to disabl13r) was created as a dichotomous variable coded '0/1'.

---

6 Form and items within forms were not fully coded largely because there were a small number of cases for these forms or items. In addition, where there was too much heterogeneity in the information recorded there was no value in coding because each code may reflect just one case.

7 Since this work was done, PDCS have introduced a more detailed set of disability codes, but given the need to collapse them to broad groups this is unlikely to have made much difference to our findings.

The same approach to coding and grouping was used to code up any 'conditions' recorded by health professionals on 'additional evidence' forms (GPFR, HFR and DBD520) present in claim packs. The variable names used on these forms consisted of form prefix (gp_, hr_ or E_) followed by disab1r to disab13r. So for example if a customer was recorded as blind on their GPFR form, the coded data for that customer would contain the variable gp_disab2r coded as a '1'.

Not surprisingly, claims mentioning malignant disease were the most likely to attract a 'correct' award decision for both the Care and Mobility components of DLA. They had the highest proportion of 'correct' decisions. Claims mentioning skin disease also had a majority of 'correct' decisions in terms of the Mobility award, but had a lower proportion of 'correct' decisions for Care awards.

Table 2.4 Grouping of conditions

Variable name (Claim form)	Grouped condition	Condition recorded by customer or health professional if on GPFR/HFR or DBD520 forms
cf_q14cond1r	Musculo-skeletal (including trauma to limbs)	Arthritis, spondylosis, back pain, disease to the muscles, bones or joints, trauma to limbs
cf_q14cond2r	Blindness/deafness (senses)	Blindness, deafness
cf_q14cond3r	Chest and breathing	Chest disease, asthma
cf_q14cond4r	Heart and blood pressure	Heart disease, cerebrovascular disease, peripheral vascular disease
cf_q14cond5r	Epilepsy	Epilepsy
cf_q14cond6r	Neurological diseases	Neurological diseases – not specific, multiple sclerosis, Parkinson's disease, chronic fatique
cf_q14cond7r	Diabetes	Diabetes mellitus
cf_q14cond8r	Mental health and behavioural disorders	Learning disability, psychosis, psychoneurosis, personality disorder, dementia, behavioural disorder, alcohol abuse, hyperkinetic syndrome
cf_q14cond9r	Disease affecting kidneys, bowel or stomach	Renal disorder, inflammatory bowel disease, bowel and stomach disease
cf_q14cond10r	Blood disorders/AIDS and multi system disorders	Blood disorders, AIDS, multi system disorders
cf_q14cond11r	Skin disease	Skin disease
cf_q14cond12r	Malignant disease (cancers)	Malignant disease
cf_q14cond13r	All others	Other answers, traumatic paraplegia/tetraplegia, major trauma

Conditions with a low proportion of 'correct' decisions included epilepsy, neurological diseases, mental health disorders and conditions affecting chest, breathing, heart and blood pressure.

Overall, the majority of decisions were the same (based on the coded data and in the real world) for both the Care and Mobility rates of DLA when customers had cancer, were blind and/or deaf or were suffering with conditions which affected their chest and/or breathing. Claims reporting neurological diseases were more likely to be assessed the same, but only for the Care component, while DMs were more likely to make the same Mobility rate assessment when the customer was suffering with diabetes.

Caution is required when interpreting associations between conditions and decisions. Customers are likely to have multiple conditions – the overall average number of reported conditions is three – and this makes it almost impossible to identify which condition causes problems for DMs. Unsurprisingly, the greater the number of conditions, the greater the number of pieces of additional evidence requested.

3 Estimation of the probability of award

This work uses the coded data from 622 claim packs. The purpose is to demonstrate the ability of a multivariate model to predict whether customers receive DLA care and mobility awards.

3.1 Data set up

As part of the data set up we derived variables based on the Department for Work and Pensions (DWP) eligibility criteria which are recorded in Appendix 1 in Kasparova *et al.* (2007). The entitlement conditions for both DLA Care and DLA Mobility were translated into derived variables using data on the claim form alone. This is set out in the remainder of this section.

It is worth noting that deriving these variables will not precisely determine eligibility because the eligibility criteria are only a guide for DMs. Eligibility is much more complicated as demonstrated by the complexity of the claim form and supporting evidence forms. If the criteria outlined below were all that were considered when deciding a DLA claim, then the claim form would only need to ask a small number of questions about needs and frequency of needs.

Given this, we have no expectation that by deriving these variables we will be able to perfectly predict DLA awards, but they are an important element of eligibility and we wanted to see how well we could predict claim awards by just looking at these variables and then seeing what difference additional evidence makes to these predictions.

3.1.1 Deriving eligibility for a care award

DLA Care eligibility was derived in several stages. First of all six components were constructed by careful translation of the DWP conditions using coded variables of customers' responses on their claims forms (see Table 3.1).

For example the first component (test1) captures the DWP eligibility condition that the customer requires 'frequent attention throughout the day in connection with bodily functions'. The coded claim form records whether the customer needs help with their toilet needs during the day and how often. The customer can also record any help required during the day with washing, bathing or looking after appearance and how often they need this help. Using the coded responses from just these two sections of the DLA claim form we derived a variable test1 if the customer needed help three or more times a day with toilet needs or help with washing/bathing two or more times a day. The threshold for the distribution of frequency of need was judged initially by looking at the frequency of need by the outcome of the claim. Subsequently, we checked the Decision Maker's Guide[8] for a definition of 'frequent' and found useful definitions which fitted well with our data findings.

The remaining five components set out in Table 3.1 were derived in a similar way. We were very careful not to over generalise a piece of data and to only use a variable where it was truly equivalent. For example, component test2 indicates whether or not the customer requires 'continual supervision throughout the day in order to avoid substantial danger to themselves or others'. The claim form asks 'Do you need supervision from another person?', 'How many days a week do you need this help?', and 'How long can you be safely left for at a time?' We define continual supervision as needing daily supervision and that they cannot be left safely for more than ten minutes.

[8] Chapter 61 of the online Decision Maker's Guide [http://www.dwp.gov.uk/docs/dmgch61.pdf]

Table 3.1 DWP definitions for eligibility to Care component of DLA with associated Public Sector Information (PSI) derived variables for capturing eligibility from coded claim form

Derived variable	DWP definitions. The customer must be so severely disabled physically or mentally that they require:	PSI description from claim form care needs variables used in derivation	Claim form coded variables
test1 (During the day)	Frequent attention throughout the day in connection with bodily functions.	Help with toilet needs 3+ times a day **or** help with washing, bathing or looking after appearance 2+ times a day.	care_needs2, care_needs3
test2 (During the day)	Continual supervision throughout the day in order to avoid substantial danger to themselves or others.	Needs supervision from another person daily and cannot be left safely for more than ten minutes.	care_needs10
test3 (At night)	Prolonged or repeated attention in connection with their bodily function at night.	Help with toilet needs (be up at least three times a night or for at least 20 minutes) every night of the week.	care_needs13
test4 (At night)	In order to avoid substantial danger to themselves or others they require another person to be awake for a prolonged period or at frequent intervals for the purpose of watching over them.	Need someone to watch over you and (be up at least three times a night or for at least 20 minutes) every night of the week.	care_needs15
test5 (Part-time care)	In connection with their bodily functions attention from another person for a significant portion of the day (whether during a single period or a number of periods).	Help with toilet needs for at least 30 minutes during the day (on average three times) or Help with washing etc for 30 minutes (twice a day).	care_needs2, care_needs3
test6 (Part-time care)	To have difficulty preparing a cooked meal for themselves if they have the ingredients.	Would you have difficulty preparing and cooking a main meal for yourself?	care_needs11

Using the derived care needs components test1 to test6 and the DWP algorithm for awarding higher, middle or lower rate DLA Care a single outcome variable was derived categorising the likely DLA Care rate awarded to a customer (see Table 3.2).

Customers were coded as likely recipients of higher rate if they satisfied either or both of the day tests (test1, test2) **and** either or both of the night tests (test3, test4). They were coded as likely recipients of middle rate if they satisfied either or both of the day tests (test1, test2) **or** either or both of the night tests (test3, test4). Finally, customers who satisfied either or both of the part-time day care tests (test5, test6) were coded as likely to receive lower rate DLA Care and customers who didn't get a higher, middle or lower rate were coded as disallowed.

Table 3.2 Algorithm of granting DLA care award

Derived variable	DWP algorithm. Customers need to satisfy:
testcare1 (high)	Either or both day tests (test1, test2) **and** either or both of the night time tests (test3, test4)
testcare2 (medium)	Either or both day tests (test1, test2) **or** either or both of the night time tests (test3, test4)
testcare3 (low)	Either or both part-time day care tests (test5, test6)
testcare0 (disallowed)	None of above satisfied

Since customers could feasibly be coded as likely to receive all three rates using this algorithm, the derivation prioritised the highest rate a customer was eligible for. The result was a singly coded variable with mutually exclusive categories for comparison with real coded outcomes (from ENT1/ DBD43 or DBD810) and for use in the modelling.

3.1.2 Deriving eligibility for a mobility award

The DLA Mobility eligibility variable was also created in several stages using coded variables from the Claim Form. Section 3 (q25 to 37) on the claim form is called 'Getting around outdoors' and records a customer's ability to walk, mobility restrictions and way of walking. The following mobility components were derived using the coded responses from the claim form (see Table 3.3 and Table 3.4).

Table 3.3 Derived mobility needs variables

Section 3 Getting around outdoors **Wording on claim form**	**Response coded by PSI**	**PSI derived components**
(q25) Are you able to walk?	No	walk1 – Unable to walk
(q26) Do you have physical problems that restrict your walking?	Yes	walk2 – Restricted walking
(q29) Walking speed	Very slow	walk3 – Very slow walking
(q30) Way you walk	Extremely poor	walk4 – Extremely poor way of walking
(q31) Do you need physical support from another person to help you to walk?	Yes	walk5 – Cannot walk without physical support
(q32) Do you stumble outdoors?	Yes	walk6 – Stumbles outdoors
(q34) Do you need someone with you to guide or supervise you when walking outdoors in unfamiliar places?	Yes	walk7 – Needs help/supervision walking outdoors
(q27) How far can you normally walk before you feel severe discomfort?	Less than 100	walk8 – Can only walk less than 100m before feeling discomfort
Customer is unable to walk or has restricted walking which is 'very slow' or 'extremely poor', requires support or is uncomfortable after 100m	Walk1=1 or (walk2=1 and (walk3=1 or walk4=1 or walk5=1 or walk8=1)	nowalk=1 – unable or virtually unable to walk
Customer is able to walk but reports some restriction when walking	Slow or very slow, poor or extremely poor, walk5=1 or walk6=1* or walk7=1 or walk8=1)	nowalk=2 – customer has some restriction when walking
		nowalk=0 – customer is able to walk without restrictions included above.

Using these eight 'mobility' components and the summary mobility variable 'nowalk' a single outcome variable was derived categorising the likely DLA Mobility rate awarded to a customer.

Customers were coded as likely recipients of higher rate mobility DLA if they were 'unable to walk' or had 'restricted walking' which was described as either 'very slow' or 'extremely poor'. Customers who reported restricted walking and required support when walking or couldn't walk further than 100m without discomfort were also allocated to the higher rate category. All other customers (those able to walk or reporting some restriction when walking) who also reported needing help or supervision when walking outdoors (walk7) were coded as likely to receive lower rate DLA Mobility. Customers who didn't get a higher or lower rate were coded as disallowed.

The low rate DWP definition was extremely hard to translate from coded variables. It includes customers with wide ranging mobility restrictions and some without any restrictions at all but who need supervision when outdoors.[9]

Table 3.4 DWP definitions for eligibility to Mobility component of DLA with associated PSI derived variables for capturing eligibility from coded claim form

Derived variable	DWP definitions. To qualify the customer:	PSI description from claim form care needs variables used in derivation	Claim form coded variables
testmob1 (high)	Must be unable or virtually unable to walk	Customer is unable to walk or has restricted walking which is 'very slow' or 'extremely poor', requires support or is uncomfortable after 100m	(nowalk=1)
testmob2 (low)	Must be so severely disabled physically or mentally that, disregarding any ability (they) may have to use routes which are familiar to them on their own, they cannot take advantage of the faculty out of doors without guidance or supervision from another person most of the time	Not (unable to or virtually unable to walk) but needs supervision when outdoors	nowalk=0 or 2, walk7=1
testmob0 (Disallowed)	Not defined	None of above	

[9] This links with the finding (not shown in this note) that customers suffering from mental illness are more likely to be awarded higher rate care/lower rate mobility. Thirteen per cent of customers suffering with psychosis, psychoneurosis or personality disorders were awarded higher rate care/lower rate mobility compared with only two per cent of customers with other disabilities/illnesses. Focusing on mobility, over half of mental health cases (53 per cent) were awarded lower rate mobility as part of their DLA award compared with only 15 per cent of customers with other types of disabilities/illnesses. In addition, customers suffering from mental health condition (psychosis, psychoneurosis or personality disorders) were found to be more likely in need of help with their medication (day and night), require supervision (outside, day and night) and need help with communication. However, it looks as though it is the 'need of help with medication' that is driving this outcome since those customers who are NOT mental health cases but who also report needing 'help/supervision with their medication' also receive higher rate DLA care/lower rate DLA mobility. This suggests that it isn't the mental health condition, but the fact that condition requires a customer to have help with, that is important. Therefore, our data confirms the views of DMs expressed at previous stages of this research, that their decisions are based on needs and not on conditions. It is also important to bear in mind that evidence for these needs may not be present on the claim form but in additional evidence.

3.2 Descriptive analysis

We use these two derived variables to compare with real world mobility and care awards.

3.2.1 DLA Care

First looking at the care awards, Table 3.5 shows the distribution of awards based on the 'real' outcome and our derived variable. Here our derived variable underestimates disallowances and higher rate care awards and overestimates lower rate care awards.

Table 3.5 Real and Derived DLA Care award

	'Real' Care award		Derived DLA Care award	
	N	%	N	%
Disallowed	203	33	147	24
Lower Rate	155	25	247	40
Middle Rate	156	25	176	28
Higher Rate	108	17	52	8
Total	*622*	*100*	*622*	*100*

Comparisons of our derived variable with the outcome of 'real' claims are shown in Tables 3.6 to 3.9. First, Table 3.6 shows whether claims were allowed or not. Looking down the columns we find that 57 per cent of claims that our derived variable suggested should be disallowed were indeed disallowed and 75 per cent of claims that our derived variable indicated should be allowed were allowed.

Table 3.6 Comparison of 'Real' awards and our Derived Care award variable

'Real' Care Award	Derived DLA Care award (column percentages)		
	Disallowed	Allowed	Total
Disallowed	84 (57)	119 (25)	203
Allowed	63 (43)	356 (75)	419
Total	*147*	*475*	*622*

Looking at the same information in a different way by considering the information in each row (see Table 3.7) we find that 41 per cent of disallowed claims would be classified as disallowed using our derived variable, whilst 85 per cent of allowed claims would be classified as allowed using our derived variable.

Table 3.7 Comparison of 'Real' awards and our Derived Care award variable

'Real' Care award	Derived DLA Care award (row percentages)		
	Disallowed	**Allowed**	**Total**
Disallowed	84 (41)	119 (59)	203
Allowed	63 (15)	356 (85)	419
Total	*147*	*475*	*622*

If we add up the cases where our derived variable correctly predicts the 'real' outcome we have 440 (84+356) out of 622 (71 per cent) correct predictions. Note that with an overall disallowance rate of 50 per cent, this prediction rate is a big improvement over chance. By contrast, if 75 per cent of cases were allowed, then predicting all cases would be allowed would be right in 75 per cent of cases.

We can then extend this analysis by looking at the rate of award for allowed claims. Looking down the columns of Table 3.8 we see that 37 per cent of claims that our derived variable indicated should receive a higher rate care award did received a higher rate care award and for middle rate and lower rate awards the figures were 31 per cent and 32 per cent. For all three award rates roughly a quarter of customers who our derived variable indicated they would receive an award were actually disallowed.

Table 3.8 Comparison of 'Real' awards and our Derived Care award variable

'Real' Care award	Derived DLA Care award (column percentages)				
	Disallowed	**Higher**	**Middle**	**Lower**	**Total**
Disallowed	84 (57)	12 (23)	45 (26)	62 (25)	203
Higher rate	30 (20)	19 (37)	28 (16)	31 (13)	108
Middle rate	20 (14)	6 (12)	55 (31)	75 (30)	156
Lower rate	13 (9)	15 (29)	48 (27)	79 (32)	155
Total	*147*	*52*	*176*	*247*	*622*

Looking at the same information in a different way by comparing row percentages (see Table 3.9) we see that 18 per cent of high rate awards would be classified as high rate using our derived variable and for middle rate and low rate awards the figures were 35 per cent and 51 per cent.

Table 3.9 Comparison of 'Real' awards and our Derived Care award variable

'Real' Care award	Derived DLA Care award (row percentages)				
	Disallowed	**Higher**	**Middle**	**Lower**	**Total**
Disallowed	84 (41	12 (6)	45 (22)	62 (31)	203
Higher rate	30 (28)	19 (18)	28 (26)	31 (29)	108
Middle rate	20 (13)	6 (4)	55 (35)	75 (48)	156
Lower rate	13 (8)	15 (10)	48 (31)	79 (51)	155
Total	*147*	*52*	*176*	*247*	*622*

Here, if we add up the cases where our derived variable correctly predicts the 'real' outcome we have 237 (84+19+55+79) out of 622 (38 per cent) correct predictions.

Further examination of mismatches reveals that, for example, out of the 203 cases which were disallowed by DMs, 62 were calculated to receive the care award at the lower rate, 45 to receive the medium rate and 12 to receive the high rate. Of the miss-matched 62 lower rate cases, 52 customers reported information on their claim form that made them satisfy either of the two part time tests with ten satisfying both. Thus, they should have been eligible for lower rate DLA care on this basis. However, DMs might have used additional information to make their final decision to disallow these particular claims.

Analysis of the additional evidence contained in these claim packs showed the miss-matches to be associated with a higher than average incidence of certain pieces of additional evidence. For example, among the miss-matched middle rate cases (n=45), 72 per cent of the packs contained a GPFR (compared with 48 per cent in the total number of packs), 35 per cent included copies of prescriptions (21 per cent in total), 14 per cent had an IB85 (nine per cent in total) and 26 per cent requested additional advice/evidence on a DBD385 form compared with only 17 per cent in the overall data. Similarly, high incidence of the need for additional evidence was found for the mismatched higher rate and lower rate cases.

Out of 108 cases awarded higher rate care, 30 should have been disallowed had DMs took information submitted on the claim pack at face value. But in this case, the miss-match can be explained by the very high incidence of DBD520 and DS1500 forms (90 per cent and 67 per cent respectively) found among these packs. These forms indicate that the customer is suffering from a terminal illness and the claim is covered under Special Rules.

The mismatches between derived and actual middle and lower rates of DLA care may be due to there being a fine line between awarding a customer the lower rate as opposed to the middle rate. Investigation of these miss-matches shows that there is certainly a difference in the amount of medical re-examination sought by DMs in these cases.

3.2.2 DLA Mobility

Table 3.10 shows the distribution of awards based on the 'real' outcome and our derived variable for mobility awards. The sample was selected to have an even distribution of awards, but this is not evident from our derived variable, where we underestimate disallowances and overestimate high rate awards.

Table 3.10 Real and Derived DLA Mobility award

	'Real' Mobility award		Derived DLA Mobility award	
	N	**%**	**N**	**%**
Disallowed	210	34	77	12
Lower rate	199	32	177	28
Higher rate	213	34	368	59
Total	*622*	*100*	*622*	*100*

We can then compare our derived variable with the outcome of 'real' claims and this is shown in Tables 3.11 to 3.14. First, Table 3.11 shows whether claims were allowed or not. Looking down the columns we find that more than three-quarters (77 per cent) of claims that our derived variable suggested should be disallowed were indeed disallowed and 72 per cent of claims that our derived variable indicated should be allowed were allowed.

Table 3.11 Comparison of 'Real' awards and our Derived Mobility award variable

'Real' Mobility award	Derived DLA Mobility award (column percentages)		
	Disallowed	**Allowed**	**Total**
Disallowed	59 (77)	151 (28)	210
Allowed	18 (23)	394 (72)	412
Total	*77*	*545*	*622*

Looking at the same information in a different way by considering the information in each row (see Table 3.12) we find that 28 per cent of disallowed claims would be classified as disallowed using our derived variable, whilst 96 per cent of allowed claims would be classified as allowed using our derived variable.

Table 3.12 Comparison of 'Real' awards and our Derived Mobility award variable

'Real' Mobility award	Derived DLA Mobility award (row percentages)		
	Disallowed	**Allowed**	**Total**
Disallowed	59 (28)	151 (72)	210
Allowed	18 (4)	394 (96)	412
Total	*77*	*545*	*622*

If we add up the cases where our derived variable correctly predicts the 'real' outcome we have 453 (59+394) out of 622 (73 per cent) correct predictions.

We can then extend this analysis by looking at the rate of award for allowed claims. Again looking down the columns of Table 3.13 we see that 71 per cent of claims that our derived variable indicated should receive a lower rate mobility award did received a lower rate mobility award and 55 per cent of claims that our derived variable indicated should receive a higher rate mobility award did received a higher rate mobility award.

Table 3.13 Comparison of 'Real' awards and our Derived Mobility award variable

'Real' Mobility award	Derived DLA Mobility award (column percentages)			
	Disallowed	**Higher**	**Lower**	**Total**
Disallowed	59 (77)	105 (29)	46 (26)	210
Higher rate	5 (6)	202 (55)	6 (3)	213
Lower rate	13 (17)	61 (17)	125 (71)	199
Total	*77*	*368*	*177*	*622*

Looking at the rows of Table 3.14 we see that 63 per cent of lower rate claims would be classified as lower rate using our derived variable and 95 per cent of higher rate claims would be classified as higher rate using our derived variable.

Table 3.14 Comparison of 'Real' awards and our Derived Mobility award variable

'Real' Mobility award	Derived DLA Mobility award (row percentages)			
	Disallowed	**Higher**	**Lower**	**Total**
Disallowed	59 (28)	105 (50)	46 (22)	210
Higher rate	5 (2)	202 (95)	6 (3)	213
Lower rate	13 (7)	61 (31)	125 (63)	199
Total	*77*	*368*	*177*	*622*

Here, if we add up the cases where our derived variable correctly predicts the 'real' outcome we have 386 (59+202+125) out of 622 (62 per cent) correct predictions.

We can then look at cases where our derived variable does not agree with the actual award. These are most often associated with the 'real' disallowances where 151 cases were expected to be allowed according to our derived variable.

Analysis of these 151 cases suggests that DMs were disregarding the responses in the claim form and taking into account (conflicting) additional evidence. Such additional evidence could have been actively sought by DMs as part of their assessment of the claim or already present in the pack. For example, 102 out of the 105 customers who our derived variable indicated that they would receive DLA mobility at the higher rate, but were actually disallowed, reported having 'restricted walking' on their claim form. Of those 102, 51 also reported having 'very slow' walking, 57 also said they 'could not walk without physical support' and 74 also reported 'needing supervision when walking outdoors'. So customers were saying all the 'right' things that could potentially lead to an award.

However, examination of the types of additional evidence present in these claim packs suggests that DMs were seeking additional evidence to the information provided on the claim forms either from customers or from health professionals and in the latter case they were over-riding customers' evidence with evidence from Health Professionals. For example, 64 per cent had a GPFR, 44 per cent had additional evidence from a DBD33 or similar, 18 per cent had an EMP and 12 per cent had an IB85. All of these were present in significantly higher proportions than in the overall claim pack data which were 48 per cent, 41 per cent, nine per cent, nine per cent respectively.

The 46 customers, who our derived variable indicated that they would receive DLA mobility at the low rate, but were actually disallowed, were also more likely to have claim packs with additional evidence.

3.2.3 Summary of descriptive analysis

To summarise, generally, the outcomes of the derived eligibility criteria variable matched extremely well with the actual outcomes of coded cases. In 71 per cent of care cases and in 73 per cent of mobility cases, the variable derived on the basis of the claim form alone could tell us whether the respective award would be granted or not. The estimates were lower (38 per cent and 62 per cent respectively) at a higher level of precision, i.e. where the rates of award were concerned. The mismatches were explained by the fact that real outcomes were based on the information from the whole claim pack and not just the claim form.

Given the relatively few coded variables employed in the derivation of our variables these results are encouraging.

3.3 Multivariate models for award receipt

We extend this analysis by using multivariate models for these predictions. Here we estimate probit models for the probability of receiving a care or mobility award and progressively increase the amount of coded evidence used in these models to see if these predictions are better with more evidence and how much better they become.

In each case the dependent variable is 1 if an award was made and 0 if not. We present four specifications of the model, run with various independent variables in each. In our first specifications reported in Tables 3.15 and 3.16 we just include the variables derived on the basis of DWP eligibility criteria discussed above. So for care there are three award rate predictions and for mobility just two. For care these three variables are 1 if the predicted award rate is lower, middle, or higher and 0 otherwise. We then add data for the health conditions reported in Table 2.4 (specification 2) and then an indicator for malignant disease from additional evidence in line with the discussion in Section 2.4 (specification 3). Our final specification (4) also includes needs based derived variables from the GPFR, DBD365G and DBD365M forms.

Using the coefficients from the models we get predicted probabilities of any award for all customers. These vary between 0 and 1, while the real outcome is either 0 or 1. Therefore, we need to define a cut off point such that any predicted probability above that cut off point we would classify as a prediction that an award would be received and a predicted probability below that cut off point predicts a disallowed claim. The cut off point we use is one that preserves the true proportion in receipt of an award.

3.3.1 Predicting whether a care award was received

The results of the model reported in column 1 of Table 3.15 come from inclusion of three indicators from our derived variables for a higher, middle or lower rate care award. All three variables are statistically significant in the model and positively associated with receipt of award. Here we get the same predictions as outlined in the previous section when we used our derived variable, i.e. 71 per cent of claims were correctly predicted.

Since this model is based on the claim form alone we sought to compare our predictions with the decisions made by DMs on the basis of coded claim forms alone. There were 81 cases where DMs said that they could make a decision based on our coding of claim form information only and reported the decisions they would make. Here we found that we correctly predicted 78 per cent of claims whilst DMs correctly predicted 83 per cent of claims.

Our predictions based on a very simple model are nearly as good as the decisions based on similar evidence made by DMs. However, this is a slightly unfair comparison because DMs were asked to predict the rate of award whilst the model is just predicted whether an award was made. It is possible that DMs would have been able to predict more accurately whether an award would be made, but were unsure as to what rate of award should be made.

Table 3.15 Probit estimates for the probability of receiving a care award

	Specifications			
	(1)	(2)	(3)	(4)
Rate according to DWP eligibility criteria, information on claim form alone	Yes	Yes	Yes	Yes
Conditions	No	Yes	Yes	Yes
Terminal illness as reported on the claim form	No	No	Yes	Yes
Needs assessments from additional evidence forms	No	No	No	Yes
Percentage correctly predicted (N-622)	71	72	73	74
Percentage correctly predicted from sample where DMs could make a decision based on coded claim form data (N-81)	78	88	89	90
Percentage correctly predicted by DMs based on coded claim form data (N-81)	83	83	83	84
Percentage correctly predicted from sample where DMs could make a decision based on all coded data (N-213)			79	78
Percentage correctly predicted by DMs based on all coded data (N-213)			80	80
Pseudo R-squared	0.06	0.13	0.15	0.17

We extended this model to include other variables from the claim form that we expect to be related to receipt of award. First adding data on the health conditions of the customers from data provided on the claim form (specification 2) we find that our percentage correctly predicted increases slightly to 72 per cent. For the sub-sample where DMs felt they could make a decision based only on claim from data our percentage of correct predictions increases substantially to 88 per cent, five percentage points higher than the correct prediction rate for the DMs for this sample.

Next (specification 3), we added a variable to the model that identifies customers with a terminal illness based on additional evidence. This increases our percentage correctly predicted again to 73 per cent and for the sub-sample where DMs felt they could make a decision based only on claim form data, our percentage of correct predictions increased slightly to 89 per cent. However, when we consider a sample based on cases where DMs could make a decision based on all our coded evidence the model correctly predicted an award for 79 per cent of cases compared with 80 per cent of correct predictions for DMs.

Finally (specification 4), we added indicators of needs assessments based on additional evidence. Again, this slightly increased our percentage of correctly predicted cases to 74 per cent and for the sub-sample where DMs felt they could make a decision based only on claim form data, our percentage of correct predictions also increased slightly to 90 per cent. Meanwhile, for the sample based on cases where DMs could make a decision based on all our coded evidence the model correctly predicted slightly fewer cases at 78 per cent of cases compared with 80 per cent of correct predictions for DMs.

3.3.2 Predicting whether a mobility award was received

The results for the same models for mobility awards are shown in Table 3.16. Using the simple specification with just the derived variables based on the mobility eligibility criteria in the model we get the same predictions as outlined in the previous section when we used our derived variable, i.e. 73 per cent of claims were correctly predicted.

For the sub-sample of 84 claims, where DMs said that they could make a decision based on our coding of claim form information only and reported the decisions they would make, we found that we correctly predicted 77 per cent of claims whilst DMs correctly predicted 83 per cent of claims.

Extending the model to include data on the conditions of the customers, and a terminal illness indicator reduces the percentage of correct predictions to 68 per cent (specification 3). It also reduces the percentage of correct predictions for the subsample of claims that DMs felt that they could make a decision about based on coded claim from data only to 77 per cent and 74 per cent. Furthermore, when looking at the sample of cases where DMs could make a decision based on all our coded evidence, the model correctly predicts 67 per cent of cases compared with 82 per cent correct predictions by the DMs.

Finally (specification 4), adding needs indicators based on additional evidence improves the percentage of correct predictions for all the samples. Overall, it increases correct predictions to 70 per cent, which is still below the percentage of correct predictions just based on our derived variable identifying the mobility entitlement criteria (specification 1). However, for the sub-sample where DMs felt they could make a decision based only on claim form data, the percentage of correct predictions from the model increased to 80 per cent, above the percentage of correct predictions from specification 1.

For the sample based on cases where DMs could make a decision based on all our coded evidence the model correctly predicted 72 per cent of cases.

It is notable that the percentage of correct predictions by DMs is higher than from all of our model specifications.

Table 3.16 Probit estimates for the probability of receiving a mobility award

Rate according to DWP eligibility criteria, information on claim form alone	Specification			
	(1)	**(2)**	**(3)**	**(4)**
Rate according to DWP eligibility criteria, information on claim form alone	Yes	Yes	Yes	Yes
Conditions	No	Yes	Yes	Yes
Terminal illness as reported on the claim form	No	No	Yes	Yes
Needs assessments from additional evidence forms	No	No	No	Yes
Percentage correctly predicted (N-622)	73	68	68	70
Percentage correctly predicted from sample where DMs could make a decision (N-84)	77	75	74	80
Percentage correctly predicted by DMs (N-84)	83	83	83	83
Percentage correctly predicted from sample where DMs could make a decision (N-215)			67	72
Percentage correctly predicted by DMs (N-215)			82	82
Pseudo R-squared	0.09	0.12	0.13	0.17

4 Conclusions

This note presented the results from estimating the probability of receiving DLA Care and Mobility awards and compared predictions from these models with predictions from DMs based on similar coded evidence sets.

The results indicated that high proportions of care (71 per cent) and mobility (73 per cent) awards can be predicted correctly using derived variables which use data on the DLA claim form and translate responses into items that reflect the entitlement criteria for care and mobility awards.

DMs were only able to make decisions on a small number of cases (less than 15 per cent) based on this information. When they did make a decision, predictions from the models were correct for only a slightly lower percentage of cases than predictions by DMs.

When additional evidence was included in the models the number of correct predictions increased slightly in the case of care awards (from 71 per cent to 74 per cent), but fell slightly in the case of mobility awards (from 73 per cent to 70 per cent). However, for the sample of cases where DMs would make decisions based on coded claim form data only, the model could correctly predict care awards in 90 per cent of cases and mobility awards in 80 per cent of cases.

The fact that inclusion of additional evidence does not improve the success of model predictions by very much, may simply reflect the fact that these are complex cases and coded forms are unable to provide enough information to make a better judgement on these cases. DMs would not make an assessment based on coded data in nearly two-thirds of cases, highlighting the complexity of many of these cases. However, our models performed reasonably well, based on relatively little information, indicating that they can be useful predictors of DLA receipt for a majority of straightforward cases.

Our earlier work highlighted that it would be possible to collect the data used in the predictive model through a bespoke survey and this could be used to estimate take-up (Kasparova *et al*., 2007). However, it is important to note that it is possible that the quarter of claims that were predicted incorrectly from the model, may have been incorrect because there were variables that predict the success of DLA claims that were not included in the model. Hence, if this were to be the case, successful estimation of DLA take-up using this approach would not be guaranteed.

References

Kasparova, D., Marsh, A. and Wilkinson, D. (2007). *The take-up rate of Disability Living Allowance and Attendance Allowance: Feasibility study*. DWP Research Report No. 442.

Purdam, K., Thornton, P., Olsen, W. K. and Afkhami, R. (2008). *Disability in the UK: Measuring Equality*. Disability and Society 23, 1 (2008): 53-65.

This working paper considers how well coded data from Disability Living Allowance (DLA) adult claim packs can predict the outcome of these claims. The probability of receiving a care or mobility award is estimated. A range of control variables are considered. First using items that reflect the entitlement criteria for care and mobility awards and then looking at whether additional evidence was required as part of the claim process. Predicted results from these estimates were then compared with real world decisions and those made by Decision Makers on the basis of the coded information from claim packs.

If you would like to know more about DWP research, please contact:
Kate Callow, Commercial Support and Knowledge Management Team,
Upper Ground Floor, Steel City House, West Street, Sheffield, S1 2GQ.
http://research.dwp.gov.uk/asd/asd5/rrs-index.asp

DWP Department for Work and Pensions

Published by the
Department for Work and Pensions
March 2011
www.dwp.gov.uk
Working Paper no. 99
ISBN 978-1-84712-939-0