Reliability of needs assessments in the community care of older people: impact of the single assessment process in England

P. Clarkson, M. Abendstern, C. Sutcliffe, J. Hughes, D. Challis

Personal Social Services Research Unit, Faculty of Medical and Human Sciences, First Floor, Dover Street Building, University of Manchester, Manchester M13 9PL, UK
Address correspondence to Paul Clarkson, E-mail: paul.clarkson@manchester.ac.uk

ABSTRACT

Background  The single assessment process (SAP) for older people, introduced in England across health and social care agencies from April 2004, aimed at improving assessment processes. We examined the impact of this policy in terms of the reliability of needs identification within statutory social services assessments.

Methods  An observational study compared the accuracy of needs identification in samples of older people before and after SAP introduction. Participants, at risk of entering care homes, were interviewed using standardized measures. Needs elicited from interviews were compared with those from statutory social services assessments to ascertain the reliability of needs identification at both times. Inter-rater reliabilities were calculated using the kappa ($k$) statistic. A Chi-squared statistic tested the equality of kappa values pre- and post-SAP.

Results  Most needs were identified more reliably after SAP introduction (range adjusted $k = 0.05–0.58$) than before (range adjusted $k = -0.09$ to 0.28), this being statistically significant for 9 out of 15 needs. Depression, and associated apathy, was an exception.

Conclusion  Statutory social services assessments better captured need following the introduction of the SAP. However, the extent to which these findings can be attributed to SAP introduction is limited by the introduction of multiple policy initiatives throughout the study period.

Background

Better assessment of older people’s health and social care needs in community settings has emerged as a challenge to improving quality of care to this group. Successive studies have identified problems in the assessment process, including evidence of unreported needs of older people; marked variability in social services assessments with neglect of information on areas such as cognitive impairment and mood state; disagreements between clinicians, nurses and social workers on the presence of particular needs and the lack of requisite health care inputs in social services assessments. These difficulties reflect a lack of consistency in how assessments are conducted and concern from the patient’s perspective over multiple contacts by different professionals involving duplication of questions and a lack of co-ordination between agencies.

One response to these problems has been a national policy introducing shared tools and procedures aimed at standardizing the assessment process. The National Service Framework in England set out several proposals, including the single assessment process (SAP), aimed at bringing together the disparate assessments of multiple professionals, greater uniformity and more effective information sharing to improve the assessment process for older people. The policy was formally introduced in England across health and social care agencies from 1 April 2004. The SAP represents an attempt to build on many of the benefits of multidimensional geriatric assessment programmes and whilst its value has been questioned this has mostly been in terms of implementation. To date, no studies have examined the impact of the SAP in settings where its changes were expected to be evident. We undertook an observational study within one English social services area to test the impact of
the policy, in terms of the reliability with which statutory community care assessments identified older people’s needs.

**Methods**

**Study population**

The study was conducted within Manchester City Council Adult Social Care, a large social services authority serving a relatively deprived population. Like all local authorities, health and primary care trusts in England, Manchester was expected to have processes in place to support the SAP from 1 April 2004. These processes included: expectations of closer professional working and shared recording procedures between health and social care, new assessment tools and a shared structure to the assessment process for health and social care agencies around four types of assessment—‘contact’, ‘overview’, ‘specialist’ and ‘comprehensive’. Manchester’s approach included the provision of training, a greater focus on sharing information with health colleagues and the use of a locally developed structured assessment tool. In the latter respect, its approach was similar to the majority (66%) of local authorities. Assessment tools in localities were often in the process of development when the SAP was being implemented and locally developed tools could be employed if they met certain standards in addition to nationally accredited ones, such as EasyCare and Functional Analysis of Care Environments. The content of Manchester’s tool concentrated on background, needs/problems identified by the service user, presence of informal carer, pattern of current support, social network, medical conditions, housing needs, activities of daily living, financial needs, psychological needs and cultural and religious background.

The study involved staff responsible for community care assessments, which required co-ordinating the work of multiple professionals, as this was likely to permit valid comparisons of the impact of the policy, particularly regarding the effects of enhanced multidisciplinary working between social workers, clinicians and nurses and the use of a more structured assessment tool. Six teams of care managers (predominantly social workers, responsible for assessments of older people under community care legislation) took part and referred participants. Manchester was chosen because comparative data on the existence, and identification, of a range of older people’s needs in the area were already available from before SAP implementation. These data were from a randomized trial of older people, the findings of which have already been reported. This study, with data collected between July 1998 and November 2000, provided a baseline (Time 1 data) from which to compare the effects of changing assessment processes. Data relating to a second period were collected between February and July 2006, between 21 and 27 months after formal introduction of the SAP (subsequently referred to as Time 2). The study objective was to determine whether older people’s needs were more successfully identified after SAP implementation than before.

All older people participating in the study were interviewed using standardized measures eliciting a range of needs considered important in the multidisciplinary assessment of older people, concentrating on those domains previously neglected by community care assessments. Interviews took place in older people’s homes, those of carers or other suitable establishments and were completed at both time periods. Researchers conducting interviews were independent of the social services personnel completing assessments.

We calculated sample sizes to establish the likelihood of achieving a statistically significant result from comparing the samples at the two time points. Sample sizes are difficult to estimate as there is little evidence of the extent to which needs are identified in this population from which to compare ratings. Given previous reports of low accuracy in identifying some needs in this setting, a relatively conservative estimate was used. We used sample size calculations for comparing two or more coefficients of reliability, or accuracy, in needs identification between the two periods. To detect a significant difference, at the $P < 0.05$ level, from a coefficient of 0.20 (‘slight’ agreement) to a coefficient of 0.60 (‘moderate’ agreement), assuming the prevalence of the need in question to be 0.50, required a sample size of 83 older people in each group at a power of 80%.

**Procedures**

We compared needs identified by care managers in their assessments with ratings of these needs by independent standardized interview at each of the two time periods. Eligible participants were those over 60 years and resident within the City of Manchester recruited as referrals from the care management teams. In each period, care managers identified older people, either recent referrals or existing cases, where a community care assessment had been completed. These cases were ones assessed for substantial levels of care or judged to have complex needs, with active consideration being given to care home admission. Thus, all older people referred to the study comprised an at-risk group and therefore were included on the basis of similar judgements regarding complexity or case-mix.

The referral process was subject to inclusion criteria, important to maintain a consistency of response from care
managers in the type of case referred. These criteria were devised to encourage suitable referrals in terms of recent health or care needs that would comprise a more 'comprehensive' assessment under the SAP. A full description of the criteria has been published.

For the standardized needs rating, each older person was interviewed by a member of the research team, usually after liaison with the care manager responsible and any carer or family member. The interviewers used a range of standardized measures designed to elicit the existence, or not, of needs independent of the judgements of care managers. The measures were: the standardized mini-mental state examination (MMSE) for cognitive impairment; the 15-item geriatric depression scale (GDS); the Barthel index for physical function; and the Clifton Assessment Procedures for the Elderly behaviour rating scale (CAPE BRS) for physical function and also apathy, communication and behaviour. Ratings of the existence of needs from these measures acted as independent gold standard judgements of the true presence of needs at the two time periods. Particular areas of need were considered in different ways. For physical functioning and behaviour, separate items from the Barthel index and CAPE BRS were scored dichotomously as the presence/absence of a difficulty. For cognitive impairment, two cut-off scores on the total MMSE score were used to define a probable case. First, a cut-off score of $<24$ was used to define a case of general cognitive impairment, traditionally used in clinical research. Second, a cut-off score of $\leq 17$ was used to denote a probable case of moderate–severe dementia; a more stringent case definition reflecting the level of cognitive impairment at which support in daily living activities becomes necessary. Similarly, for depression, a cut-off score on the GDS of $>5$ was used, reflecting a case definition of probable depression.

To obtain information on needs identified by care managers, researchers screened the assessment documents relating to each referred case within the agency. Data were extracted from the assessment document and relevant supporting information (such as correspondence) using a pro forma listing each relevant need, comparable to that contained in the standardized measures, and whether the need had been identified at any level of severity by the care manager (including when this had been stated after consultation with health colleagues).

**Statistical analysis**

The characteristics of the samples, at both periods, were compared using $t$-tests for interval level data and Chi-squared tests for ordinal level data. This analysis assessed the equivalence of the samples in terms of characteristics and the prevalence of needs (within-sample estimates from the standardized measures) to test the robustness of conclusions regarding the impact of other factors, such as those arising from the policy change.

Inter-rater reliabilities between social services and standardized assessments on the judgements of needs at each time period were calculated using the percentage of agreement and the kappa statistic ($k$). Reliability refers here to the consistency of ratings (whether a need exists or not) on a specific case. Kappa measures the degree of difference between observed agreement and that expected by chance (expected agreement). This difference is expressed as a score between $-1$ and $1$, where $1$ is perfect agreement, $0$ is what would be expected by chance and a negative value is agreement less than chance. However, because an unadjusted kappa value may not take account of the prevalence of need and the propensity of different observers to identify it, the maximum possible value of kappa ($k_{max}$) was also calculated using the equation: $1 - (\text{minimum disagreement/expected disagreement})$. The ratio $k/k_{max}$ was used as a general measure denoting the degree of agreement beyond chance within these constraints. This adjusted value is useful when comparing different populations, as here, when the intention is to interpret the different degrees of agreement that may result. The overall kappa coefficients at Time 1 and Time 2 were compared for each area of need using a test statistic that follows a Chi-squared distribution, with appropriate probability values therefore reportable. This analysis tested the hypothesis that, in each need domain, the accuracy of needs identification was significantly higher post-SAP implementation.

We also constructed logistic regression equations to model the effects of ‘exposure’ to the policy (Time 2 versus Time 1) on the probability of detecting each need (presence/absence), controlling for its prevalence as a potential confounding effect. This analysis was to determine the effects of the changed assessment arrangements allowing for the prevalence of each need at the two time periods, in effect controlling for the potential to recognize needs more easily if more of them are present. All analyses were done with SPSS version 14 (SPSS, Chicago, IL). We report as significant effects at the 5% level, with 95% confidence intervals (CI) reported for unadjusted kappa values.

**Results**

Over the pre-SAP period, after exclusions, there were 127 older people referred for whom a full social services assessment was available. Of these, 106 (83%) agreed to take part.
and were interviewed. Corresponding figures for the post-SAP period were 134 referrals with 110 (82%) being interviewed. The demographic characteristics of those interviewed did not differ significantly from those who refused or were unavailable for interview at both time periods ($P > 0.05$ for all variables).

Table 1 reports the characteristics of older people interviewed before and after the introduction of the SAP, showing equivalence between the samples. Older people in both periods were relatively old (a mean age of 82 years), predominantly female, white, widowed or single, living alone and in their own homes (either owned or rented).

### Table 1: Characteristics of older people pre- and post-SAP

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Time 1 (pre-SAP)</th>
<th>Time 2 (post-SAP)</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>106</td>
<td>110</td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>82 (7.7)</td>
<td>82 (7.2)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>77 (73)</td>
<td>66 (60)</td>
<td></td>
</tr>
<tr>
<td>White ethnicity</td>
<td>105 (99)</td>
<td>105 (95)</td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>22 (21)</td>
<td>24 (22)</td>
<td></td>
</tr>
<tr>
<td>Married/cohabiting</td>
<td>17 (16)</td>
<td>21 (19)</td>
<td></td>
</tr>
<tr>
<td>Widowed</td>
<td>64 (60)</td>
<td>57 (52)</td>
<td></td>
</tr>
<tr>
<td>Living alone</td>
<td>70 (66)</td>
<td>66 (60)</td>
<td></td>
</tr>
<tr>
<td>Living with spouse/partner</td>
<td>14 (13)</td>
<td>20 (18)</td>
<td></td>
</tr>
<tr>
<td>Living with others</td>
<td>22 (21)</td>
<td>24 (22)</td>
<td></td>
</tr>
<tr>
<td>Living in own home</td>
<td>69 (65)</td>
<td>82 (74)</td>
<td></td>
</tr>
<tr>
<td>In permanent accommodation</td>
<td>88 (83)</td>
<td>96 (87)</td>
<td></td>
</tr>
<tr>
<td>Prevalence of needs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Help with feeding*</td>
<td>10 (9)</td>
<td>38 (34)</td>
<td>$&lt;0.001$</td>
</tr>
<tr>
<td>Difficulty in transfers*</td>
<td>20 (19)</td>
<td>54 (49)</td>
<td>$&lt;0.001$</td>
</tr>
<tr>
<td>Difficulty grooming*</td>
<td>28 (26)</td>
<td>63 (57)</td>
<td>$&lt;0.001$</td>
</tr>
<tr>
<td>Difficulty toileting*</td>
<td>20 (19)</td>
<td>44 (40)</td>
<td>$&lt;0.001$</td>
</tr>
<tr>
<td>Difficulty bathing*</td>
<td>85 (80)</td>
<td>92 (84)</td>
<td></td>
</tr>
<tr>
<td>Help with mobility*</td>
<td>88 (83)</td>
<td>76 (69)</td>
<td>$&lt;0.05$</td>
</tr>
<tr>
<td>Difficulty with stairs*</td>
<td>65 (61)</td>
<td>87 (79)</td>
<td>$&lt;0.05$</td>
</tr>
<tr>
<td>Help with dressing*</td>
<td>42 (40)</td>
<td>73 (66)</td>
<td>$&lt;0.001$</td>
</tr>
<tr>
<td>Incontinent (bladder or bowel)*</td>
<td>30 (29)</td>
<td>47 (43)</td>
<td>$&lt;0.05$</td>
</tr>
<tr>
<td>Case moderate–severe dementia*</td>
<td>20 (19)</td>
<td>39 (37)</td>
<td>$&lt;0.05$</td>
</tr>
<tr>
<td>Probable case depression*</td>
<td>33 (33)</td>
<td>36 (35)</td>
<td></td>
</tr>
<tr>
<td>Apathy*</td>
<td>99 (94)</td>
<td>107 (97)</td>
<td></td>
</tr>
<tr>
<td>Communication difficulties*</td>
<td>10 (9)</td>
<td>25 (23)</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td>Social disturbance*</td>
<td>70 (68)</td>
<td>59 (54)</td>
<td>$&lt;0.05$</td>
</tr>
</tbody>
</table>

Data are mean (SD) or number (%).

*Response of ‘dependent/unable’ or ‘needs help’ from Barthel index.

*Response of ‘unable—no sitting balance’, ‘major help, can sit’ or ‘minor help—verbal or physical’ from Barthel index.

*Response of ‘needs help with personal care’ from Barthel index.

*Response of ‘immobile’, ‘wheelchair independent’ or ‘walks with help’ from Barthel index.

*Response of ‘incontinent’ or ‘occasional accident’ from Barthel index.

†Defined as a total score on the standardized MMSE of $\leq 17$.

‡Defined as a total score on the GDS of $>5$.

§Defined as a problem ‘sometimes’ or ‘frequently’ from sub-scales of the CAPE BRS: apathy (five statements on difficulties with supervision outside, helping out, constructive occupation, socializing and willing to do things suggested); communication (two statements on difficulties with understanding and communicating by speaking, writing or gestures); social disturbance (five statements on difficulties with objectionable behaviour day/night, accusations, hoarding and disturbed sleep).
These characteristics mirror those found in surveys of older people entering care homes, adding validity to the entry criteria for the study. However, there were differences in the prevalence of needs, with more interviewed at Time 2 having difficulties with activities of daily living, incontinence, cognitive functioning and communication.

Table 2 shows data relating to the accuracy of needs identification across the two periods. In all areas, with the exception of depression and the associated behaviour ‘apathy’, agreement between social services assessments and interview was higher at Time 2. In the areas of feeding, transfers, grooming, toileting, mobility, difficulties with stairs, incontinence and cognitive impairment (including case level dementia), this difference was statistically significant. For some areas, such as difficulties with feeding, managing stairs, and whether a case of dementia was identified, agreement at Time 1 was very low. Post-SAP implementation (Time 2), however, the adjusted kappa values suggest that there remains room for improvement with agreement being only ‘slight’ to ‘moderate’ in most areas (range 0.05–0.58). For some areas, such as depression, adjusted kappa values suggest only ‘slight’ agreement between social services assessments and the standardized measures, this remaining appreciably the same before and after SAP implementation.

Similarly, the results of logistic regression analyses showed that, in all areas of need, the probability of detecting needs was significantly increased after SAP (Time 2), with the exception of depression, apathy, communication difficulties and social disturbance (as above). Cognitive impairment, for example, was almost four times more likely to be detected after allowing for its prevalence at the two time periods (odds ratio = 3.82; 95% CI = 1.81 to 8.04; P < 0.001). The prevalence of each of these needs did not have a significant impact on detection, suggesting that the SAP had impacted positively on detection despite, and independently of, the existence of each area of need. Interaction terms were also added to the models to reflect co-morbidity. For example, examination of the impact

<table>
<thead>
<tr>
<th>Needs identified</th>
<th>Time 1 (pre-SAP)</th>
<th>Time 2 (post-SAP)</th>
<th>( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical functioning( ^b )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeding</td>
<td>71</td>
<td>-0.08 (-0.25 to 0.09)</td>
<td>-0.09</td>
</tr>
<tr>
<td>Transfers</td>
<td>69</td>
<td>0.04 (-0.15 to 0.23)</td>
<td>0.04</td>
</tr>
<tr>
<td>Grooming</td>
<td>54</td>
<td>0.005 (-0.17 to 0.18)</td>
<td>0.006</td>
</tr>
<tr>
<td>Toileting</td>
<td>65</td>
<td>0.003 (-0.19 to 0.19)</td>
<td>0.003</td>
</tr>
<tr>
<td>Bathing</td>
<td>67</td>
<td>0.10 (-0.08 to 0.29)</td>
<td>0.11</td>
</tr>
<tr>
<td>Mobility</td>
<td>66</td>
<td>0.01 (-0.17 to 0.20)</td>
<td>0.01</td>
</tr>
<tr>
<td>Difficulty with stairs</td>
<td>49</td>
<td>-0.03 (-0.21 to 0.16)</td>
<td>-0.03</td>
</tr>
<tr>
<td>Dressing</td>
<td>53</td>
<td>0.04 (-0.15 to 0.23)</td>
<td>0.04</td>
</tr>
<tr>
<td>Incontinence (bladder or bowel)</td>
<td>64</td>
<td>0.14 (-0.05 to 0.34)</td>
<td>0.14</td>
</tr>
<tr>
<td>Mental health</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive impairment( ^c )</td>
<td>44</td>
<td>0.05 (-0.11 to 0.21)</td>
<td>0.07</td>
</tr>
<tr>
<td>Case moderate–severe dementia( ^d )</td>
<td>57</td>
<td>-0.08 (-0.25 to 0.09)</td>
<td>-0.09</td>
</tr>
<tr>
<td>Depression( ^e )</td>
<td>67</td>
<td>0.22 (0.06–0.38)</td>
<td>0.28</td>
</tr>
<tr>
<td>Behaviour( ^f )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apathy</td>
<td>14</td>
<td>0.01 (-0.02 to 0.04)</td>
<td>0.07</td>
</tr>
<tr>
<td>Communication difficulties</td>
<td>90</td>
<td>0.22 (0.04 to 0.39)</td>
<td>0.23</td>
</tr>
<tr>
<td>Social disturbance</td>
<td>67</td>
<td>0.02 (-0.08 to 0.13)</td>
<td>0.04</td>
</tr>
</tbody>
</table>

\( K = \) Cohen’s kappa, unadjusted; \( k_{\text{max}} = \) kappa adjusted to reflect the maximum value it could attain.

\( ^a \)Defined as agreement on ‘cases’ and ‘non-cases’.

\( ^b \)Social services assessment of need compared with rating on Barthel index.

\( ^c \)Social services assessment of need compared with rating on standardized MMSE.

\( ^d \)Social services assessment of need compared with rating on GDS.

\( ^e \)Social services assessment of need compared with rating on CAPE BRS.

\( ^f \)Chi-squared statistic with 1 degree of freedom, after Donner et al. 46
of the SAP on the probability of detecting cognitive impairment, allowing for its prevalence and also the prevalence of cognitive impairment with depression or with incontinence. No significant interaction effects were found.

Discussion

Main findings

Improvements in the reliability of identifying older people’s needs within statutory social services assessments were found in one area after the introduction of the SAP in England. These assessments were expected to involve closer working and information sharing between health and social care professionals after the policy’s introduction. Participants were included on the basis of being at risk of entering care homes and therefore would be expected to meet the criteria for ‘comprehensive’ assessment within the SAP.

What is known already

Other studies have testified to the potential benefits of involving multiple professionals in sharing information as part of assessments at this critical point in an older person’s care. However, poor quality information was collected on cognitive impairment within community care assessments and assessments of toileting and incontinence were often cursory. Our study showed that the identification of these conditions had improved since the introduction of the SAP. One way of achieving this may have been through the closer involvement of specialist clinicians in social services decision making through their ‘specialist’ contribution to assessments as part of the SAP. Another possible mechanism for improving assessment of these conditions may have been through training in assessment, some of which was provided by health colleagues.

Previous studies examining the accuracy of needs identification in older people have shown differential effects for particular conditions. For example, occupational therapists’ ratings on Barthel items reflecting motor skills (e.g. bathing, mobility and stairs) have shown higher agreement than those reflecting cognitive functioning (e.g. bowels, bladder and grooming). This was not the case for our study, where problems reflecting cognitive difficulties showed similar or higher agreement than those reflecting motor skills, although difficulties before SAP implementation were almost universally poorly detected. It may be, as previous studies have suggested, that different professional groups have a tendency to identify different conditions or problems. One challenge for the SAP will be to enable different professionals to become aware of the existence of a range of conditions and to accept the assessments made by others.

The policy’s impact on clinical practice has been viewed negatively by some. A survey in old age psychiatry reported that referrals from general practitioners after SAP introduction took longer to read, were more illegible and contained less information. These changes were thought to have impaired communication. However, this particular study judged the policy’s impact from the viewpoint of working practices rather than, as here, in terms of the broader aim to improve the accuracy of needs identification.

What this study adds

This is the first study using patient/user level data to examine the impact of the SAP in a setting in which the benefits of closer co-operation between health and social services in an overall assessment of needs may be discerned. Need identification is the first step to effective intervention and our findings show broad improvement after SAP introduction.

However, our study showed that, despite this improvement, there remained considerable room for further gain in the identification of some needs of older people. Depression in particular was poorly identified both before and after the SAP and this is particularly relevant in the light of previous studies testifying to the poor detection of this condition across multiple settings. This finding presents a challenge to assessment policies and the use of shared tools, procedures and training in recognition may be an area for further development. The accurate detection of depression may be assisted by the use of standardized tools, often where its identification is included as a subset to a broader assessment. However, in the social services setting, the actual detection of the disorder is arguably less important than determining its likely effects on functioning and needs for care. Thus, if not properly identified, depressive symptoms may persist and may lead to increased isolation and breakdown in family relationships. Failure to detect needs arising from this, and other conditions, may therefore lead to a lack of appropriate management and necessitate more complex care packages and referral to other professionals such as specialist clinicians. This would have resource consequences as well as incur personal costs to the user. There are, therefore, important implications from unrecognized needs in terms of planning care and in potential future demand for more specialist help. However, the potential benefits of incorporating standardized scales into social services assessments, as a means of improving recognition, have to be balanced against the costs, including the financial costs of licence agreements and the opportunity costs of increased professional time taken on administration.
More generally, while the SAP did appear to improve the identification of a range of needs as part of assessments in this area, the accuracy of assessment was still unacceptably low. However, it is not only within the social services setting where this is the case. In primary care, for example, the detection of important conditions in older people, such as depression, is also very low. There is therefore clearly room for improvement in the detection of a range of conditions, an area of improvement needed not just in the context of social care.

It will be important in evaluating the policy's overall impact to examine the contribution of the different elements within the guidance to improving the assessment process. Local implementation around methods and levels of assessment therefore remains an issue for future research. While the reported gains here are likely to be attributable to the SAP preparatory and training procedures, the role of other aspects, such as the use of standardized tools, is harder to discern. There are further implications for clinical practice in the importance attached to particular areas of involvement in the policy by specialist clinicians. Linking these areas together will offer a more detailed picture of the processes through which the reported benefits described here are mediated. Further work is needed to explore these aspects and their relationship to impact.

**Limitations of this study**

It is difficult to isolate the effect of the phenomenon of interest (the SAP policy change) from other influences ongoing at the time of the study. Although the groups were similar, some needs were more prevalent after the introduction of the SAP. This may have been influenced by other policies, such as the Fair Access to Care guidance, which may have led to the application of more restrictive eligibility criteria by the social services authority. The threshold required to receive assistance may have been raised and thus older people entering the system post-SAP may have been more dependent than previously, perhaps making detection of needs more likely. Although this might have been a possible confounding effect in the study, our adjusted kappa ratings did take account of the prevalence of needs in each sample. Furthermore, regression analyses to control for this showed a negligible effect of prevalence on the identification of needs. Nonetheless, our study, by its nature, did not allow for multiple policy influences and so it is difficult to conclude unequivocally that the SAP alone was instrumental in improving the accuracy of needs identification.

The study was also conducted in only one agency where data on assessments were available before the introduction of the SAP. Caution is required in generalizing to agencies in other localities who may have implemented different policies and procedures for assessment. In this respect, the degree of variation between agencies has been widespread in terms of factors such as the use of different tools and the way SAP was integrated with other assessment procedures such as those for intermediate care. Local implementation has emerged as an issue for the policy since its introduction. One difficulty in interpreting the changes consequent on the policy has been its multiple aims, directed at resolving several problems with the assessment process from the perspectives of different professionals. This study therefore analyses the policy impact from one, albeit important, viewpoint and further work is necessary to examine the broader impact of the policy.

**Acknowledgements**

We thank all care managers in the social services teams for referring participants to the study; Caroline Marsh, Alan Calvert and Roger Lightup for sanctioning the research and data collection through Manchester City Council Adult Social Care; the older people and their carers for agreeing to participate; and Professor Graham Dunn for statistical advice.

**Funding**

The study was funded by the Department of Health in England. The views expressed here are those of the authors and are not necessarily shared by the Department of Health.

**References**

26 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.

47 skaane P, Engedal K, Skjennaldal A. Interobserver variation in the interpretation of breast imaging. Comparison of mammography,


