PSSRU

Measuring the outcomes of care homes: Final report

‘Measuring Outcomes for Public Service Users’ Project

Ann Netten, Julie Beadle-Brown, Birgit Trukeschitz, Ann-Marie Towers, Elizabeth Welch, Julien Forder, Jan Smith and Elaine Alden

---

PSSRU Discussion Paper 2696/2
June 2010
www.pssru.ac.uk

PERSONAL SOCIAL SERVICES RESEARCH UNIT

University of Kent

University of Kent
Cornwallis Building
Canterbury
Kent
CT2 7NF
Tel: 01227 823963/823862
PSSRU@kent.ac.uk

London School of Economics

London School of Economics
LSE Health & Social Care
Houghton Street
London
WC2A 2AE
Tel: 020 7955 6238
PSSRU@lse.ac.uk

University of Manchester

University of Manchester
Dover Street Building
Oxford Road
Manchester
M13 9PL
Tel: 0161 275 5250
PSSRU@manchester.ac.uk

www.PSSRU.ac.uk

Contents

1. Introduction

1.1 Background

The demand for long-term care services is predicted to rise, primarily as a result of longer life expectancy, as people with existing long-term conditions live longer and more people survive into very old age (Wanless, 2006; Emerson and Hatton, 2008; Department of Health, 2009*).* While many people with long-term impairments can continue to be productive with minimal support, caring for those with high levels of impairment incurs high costs, whether these costs are borne by the individual, their families, the third sector or the state. Inevitably, resources are limited. If we are to target our resources to incur the most benefit, we need to be able to identify, measure and monitor that benefit. This is true for central government in allocating funding across a wide range of activities; for those commissioning long-term care and other services and support; and those providing and quality assuring these services. This raises many challenges, including being able to attribute benefit or outcomes to services and support at the same time as minimising the burden to all concerned in collecting the data.

We report on a project which developed and tested an approach to measuring and monitoring outcomes of the care and support provided to residents of care homes for older people and people with learning disabilities. Care homes represent the most intensive type of support provided to some of the most disabled members of the population. The research was part of the 'Measuring Outcomes for Public Service Users' (MOPSU) project, which was funded over three years (2007-2009) by the Treasury under the Invest to Save budget and led by the Office for National Statistics (ONS). The MOPSU project consists of 3 main work-strands all working towards the overall aims of:

- more efficient and effective commissioning and procurement of services, placing the issues of quality and value for money at the heart of the decision-making process
- encouraging the use of 'outcomes' measures to assess the impact of services on their users, across the spectrum of providers
- examining the extent to which the third sector is involved in public service delivery and helping to alleviate barriers to entry to third sector organisations

Work undertaken at PSSRU focused primarily on developing approaches to measuring and monitoring the value of social care. The overarching aim was to develop the Adult Social Care Outcomes Toolkit (ASCOT) that would provide a variety of approaches to identifying and monitoring value across the range of social care interventions.

We start by outlining our overall approach to identifying the value of public services and identify the overarching aims of the study. We then discuss how we define, identify and measure outcome and the type of information that is already available, through the regulator, on which we might draw when monitoring the value of care homes.

1.2 Identifying the value of public services

We are proposing an indirect approach to measuring the value of services by which we identify the potential value that could be delivered by a service or intervention: the *Capacity for Benefit (CfB)* and the degree to which that value is actually delivered through a measure of *quality (Q)*. CfB is the potential benefit that could be delivered if all the objectives of a service or intervention were achieved for all those receiving the service. This depends on:

- The number of beneficiaries
- The domains of outcome the service aims to influence
- The level of well-being or needs in these domains in the absence of the service or intervention
- What is hoped to be achieved in each domain

In terms of what is hoped to be achieved, social care services have traditionally been and still predominantly are, concerned with meeting long term care needs or ‘maintenance’ of service users (Qureshi et al., 1998). However there is increasing policy emphasis on enablement and prevention (Ministers, Local Government et al., 2007), where we might expect to see changes in people’s abilities to meet their own needs.

Quality reflects the level of outcome actually achieved in the relevant domains and includes those aspects of the care process that form another dimension to the lives of those in receipt of long-term care, defined as ‘dignity’ (see below).

While we are particularly concerned here with social care, this basic framework has a broader applicability. One advantage of this approach is that we can identify the full impact of the intervention rather than marginal changes, which are reflected in before and after measures. Moreover, a global indicator of output can only be used to establish whether productivity is rising or falling. In this approach we can distinguish whether changes in output are associated with changes in what the intervention or service in its current configuration can deliver or whether this is due to improvements/reductions in the quality of what has been delivered.

Previous work (Netten et al., 2006) demonstrated how the approach can be used to reflect the impact on outputs of care homes for older people of increased levels of dependency and changing levels of quality reported by the regulator (then the Commission for Social Care Inspection (CSCI)). Necessarily, this was based on a number of assumptions and available data. This project builds on that work to develop a validated approach to measuring and monitoring the value of outputs of care homes.

1.3 Study aims

The overall aims of the study were to:

- Develop and validate a method of measuring and monitoring outcomes based on routinely available regulator quality ratings that would allow comparison over time, between homes and across client groups
- Develop and validate a method for monitoring CfB through the dependency of residents, in a way that can be used to reflect changing value of care home outputs over time

This was to be achieved by establishing valid and reliable measures of CfB and outcome for a sample of residents and to relate these to data that are routinely collected and reported by the regulator, or that could be.

Particular methodological challenges were presented as the majority of residents of care homes are very vulnerable, many having cognitive impairment and/or communication difficulties. In order to achieve the overall aims we needed to explore the degree to which we can reliably establish CfB and outcome for people with these types and levels of impairment in a care home setting. As part of this we were investigating the applicability of an approach that had largely been developed with older people to younger adults with learning disabilities.

There were a number of research questions around the provision of care and support to people from ethnic and cultural minorities.[1] These issues were the subject of a stream of work linked to this study but are not reported on here as they were the focus of an ESRC collaborative CASE studentship[2] for one of the authors (JS).

1.4 Defining outcomes of social care and support

In earlier conceptual work, completed as part of this project and described in our interim report (Forder et al., 2007), we identified that the main goal of social care is the improvement in *well-being* or *quality of life* that people experience in using the service, so for particular services we aim to measure the improvement in quality of life or *outcomes* they confer to service users.

In previous work we argued that outcomes reflect our fundamental motives: quality of life, well-being, happiness, utility and so on. These are hard to define specifically enough to allow a measurement framework to be developed. Instead, we break down these fundamental motivations into important domains that comprise *social care-related quality of life* (SCRQOL). A series of studies has investigated and tested the domains most relevant to social care in a number of different contexts, including the sister MOPSU project on low-level services that focused on day care (Caiels et al., 2010) and an evaluation of individual or personal budgets (Glendinning et al., 2008). Another ongoing project, which is developing a 'gold standard' preference weighted measure of outcome, has analysed data based on previous work and cognitively tested the domains listed in Box 1.1 (see Netten et al., 2009). These domains formed the basis of the measures used in this study.

---

[1] From a methodological point of view, are our measures accurately reflecting the well-being of people from these groups? Are different aspects of care more important to them than to people in the mainstream white British group? Are there aspects of need that are less well met among these groups due to lack of knowledge or discrimination? Are their cultural traditions and preferences adequately reflected in the way care is provided?

[2] The studentship was supported by CSCI, PSSRU and Tizard.

Box 1.1: Social care-related quality of life (SCRQOL) domains

Domains of SCRQOL	Definition
Personal cleanliness and comfort	The service user feels he/she is personally clean and comfortable and looks presentable or, at best, is dressed and groomed in a way that reflects his/her personal preferences
Safety	The service user feels safe and secure. This means being free from fear of abuse, falling or other physical harm and fear of being attacked or robbed
Control over daily life	The service user can choose what to do and when to do it, having control over his/her daily life and activities
Accommodation cleanliness and comfort	The service user feels their home environment, including all the rooms, is clean and comfortable
Food and nutrition	The service user feels he/she has a nutritious, varied and culturally appropriate diet with enough food and drink that he/she enjoys at regular and timely intervals
Occupation	The service user is sufficiently occupied in a range of meaningful activities whether it be formal employment, unpaid work, caring for others or leisure activities
Social participation and involvement	The service user is content with their social situation, where social situation is taken to mean the sustenance of meaningful relationships with friends, family and feeling involved or part of a community should this be important to the service user
Dignity	The negative and positive psychological impact of support and care on the service user's personal sense of significance

For the most part the SCRQOL domains described in Box 1.1 are fairly self-explanatory. However, it is worth briefly discussing the Dignity domain, which poses particular challenges in the measurement of outcome. As we identify above, this reflects an aspect of people's lives that is not relevant in the absence of care and support: the care process. All other domains are relevant whether or not we have care and support needs. We can define our outcomes in terms of the difference between our quality of life in these domains with and without care and support. The care process does not exist in the absence of care and support. When care is present our sense of self is as likely (indeed it is arguable that it is more likely) to be undermined as enhanced by the way that people supporting us behave towards and treat us.

Because of these differences, the inclusion of the Dignity domain in measuring outcome raises particular methodological challenges. As a result of this and, as we describe in Chapter 4, some concerns about reliability of ratings in this domain, we did not include Dignity in our final estimates.

In order to identify the full impact of services we identify both what current SCRQOL is and what it is expected to be in the absence of formally provided social care and support. The difference between experienced or current SCRQOL and 'expected' SCRQOL is the outcome of the service or intervention. An individual's capacity *to* benefit (CtB) from a service is the difference between his or her expected SCRQOL and the maximum level that they could achieve, given what the service is providing. An interventions' capacity *for* benefit is the mean level of service users capacity *to* benefit.

$$CfB_i = \sum_{jk} \frac{MaxO_j - EO_{jk}}{n}$$

Where:

CfB_i is the capacity for benefit for service i

$MaxO_j$ is the maximum score for outcome domain j

EO_{jk} is the expected needs for individual k in domain j in the absence of service i

n is the number of individuals for whom we have information about needs in the absence of service i.

The actual level of quality adjusted output of a given intervention is the number of people served multiplied by the interventions' CfB and quality.

$$Op_i = CfB_i \, Q_i \, N_i$$

Where:

Op_i is total quality adjusted output of service i

Q_i is quality defined as the level of SCRQOL actually achieved (current/experienced SCRQOL)

N_i is the number of people using service i.

1.5 Measuring SCRQOL and outcome

In order to create measures of capacity to benefit (CtB) and outcome (defined here as the impact of the service) we need basic building blocks for each of our domains of current SCRQOL (or need levels) and expected SCRQOL (or expected need levels in the absence of services). With these we can derive (i) CtB from the difference between the expected and maximum level of SCRQOL and (ii) SCRQOL gain, or service outcome, from the difference between expected and current, or experienced, SCRQOL.

A basic principle underlying the identification of SCRQOL and outcomes is that they should reflect as far as possible the perspective of the service user. The definitions of each domain, as given in Box 1.1, are described in terms of how people feel. We are more concerned with whether people *feel* safe than whether they are in some ‘objective’ way free from harm. Elsewhere (Forder et al., 2007; Netten et al., 2009) we distinguished between ‘functioning’ – how well people function in these domains, and what Sen terms ‘capabilities’ – people’s potential to function at any given level (Sen, 1985; Sen, 1993). For example, we could define an individual’s social participation and involvement by how often they see people (functioning) or, reflecting the fact that people vary in how often they want to see people, we could define it in terms of how happy they are with their level of social contact (capability).

In the low level services (LLS) project (Caiels et al., 2010), the objective was to reflect capabilities as far as possible: how people felt about their situation. In this project and in other work (Netten et al., 2006), estimates of expected need in the absence of services were based on interviews with mentally intact older people living in their own homes. People were asked what their current level of well-being was in each domain, whether services helped them in this aspect of their lives and, if services do help them, what they would judge to be their needs or situation in the absence of that help. In this situation, people are able to reflect on what their levels of need were likely to be in the absence of services (Netten et al., 2006).

This is particularly challenging when we consider that for a substantial proportion of the social care population, their level of communication and cognitive skills severely limit them in responding to the type of structured questions used to establish people's experienced SCRQOL. As we describe in Chapter 2, our principal approach to establishing people's level of need or current SCRQOL in each domain was to gather evidence from a variety of sources: through observation, both structured observational periods and more generally; through interviews with people when they were able to respond; and through interviews with staff. Necessarily, such judgements, particularly when based just on observation and the views of staff, will be closer to measures of 'functioning' than capabilities.

When creating a composite measure of SCRQOL across all our domains, any simple additive indicator will be based on an implicit assumption that all of these domains and the levels of need within them are equal. If we are to accurately reflect the value of a service in a way that allows comparison of the benefits with other interventions, or other ways resources could have been used, we want to reflect the relative importance of these domains. We do this by weighting them to reflect preferences. Whose preferences should be used is arguable, but in the case of the use of public money the preferences of the general population provide a legitimate basis for weighting outcome (De Wit, 2000). For the purpose of this study preference weights were drawn from a previous preference study conducted for another research project (Burge et al., 2006). A further population preference study was conducted as part of the MOPSU project (Burge et al., 2010).

1.6 Regulation of care homes

As we describe above, one of the key objectives of the work is to make use of information that was already available, about the quality of care homes and nature of the population they serve, or could be made available through this route.

The nature of that information and the regulatory process has changed over recent years, resulting in a variety of ways of reporting quality. National minimum care standards were legislated for under the Care Standards Act 2000. The regulator (initially, the National Care Standards Commission (NCSC)) reported on these standards from 2002/03. Initially all homes were rated every year as whether they had 'exceeded', 'met' , 'almost met' or 'not met' all these standards. The successor to NCSC, the Commission for Social Care Inspection (CSCI), aimed to move to an outcome (rather than input) focus and introduced Inspecting for Better Lives (Commission for Social Care Inspection, 2005). This identified a subset of these standards as 'key' on which homes were rated every year. The standards were also classified into eight 'outcome groups' that homes were inspected for under the Key Lines of Regulatory Assessment (KLORA). These ratings provided the basis for the overall quality rating of the homes as zero (poor), one (adequate), two (good) or three-star (excellent). Homes were first star rated in 2007. From April 2008 the star ratings of homes were made publicly available on the CSCI (and subsequently Care Quality Commission (CQC)) website.

The regulatory process has also changed. Since 2008 homes are no longer automatically inspected at least once and usually twice a year as had been the case previously. During the period of the study the main inspection was called a 'key inspection'. Excellent homes were usually inspected only once every three years. Good homes were inspected at least once every two years, adequate homes at least annually, and poor homes usually twice a year.[3] In addition, homes may have received a random inspection to follow up issues raised in a key inspection or to check that a good or excellent home was maintaining its standards. Finally, some services might be selected for what was known as a thematic inspection: an inspection focusing on a particular issue, such as medication.

All homes were required to return an Annual Quality Assurance Assessment (AQAA) form, which includes, *inter alia,* self ratings of quality, and information about residents and staff. The inspection data used in this study were the data from the most recent key inspection.

In planning the study we were aware that, particularly given the historical emphasis on structure and process in regulation, many other factors than resident outcomes were likely to be associated with quality ratings, potentially making the relationship between these weak. Indeed, previous research (Beadle-Brown et al., 2008a) had found no relationship between the proportion of standards met and outcomes in terms of engagement among residents in homes for people with LD. Moreover, the regulatory process and basis for quality ratings was again under review at the time that this study was reporting (Care Quality

---

[3] http://www.cqc.org.uk/guidanceforprofessionals/socialcare/careproviders/inspection/keyinspection.cfm)

Commission, 2010). This had the disadvantage that linking our in depth measures of outcome to data available through the regulator at the time of the study would not achieve the overarching aim of enabling commissioners and others to use routine sources to monitor outcomes, as the ratings and data collections were in place for such a short period. However, the advantage of the timing was that the results of the study were emerging at a time to be of most use to the regulator in designing new systems, better able to reflect the experiences of and outcomes for residents.

1.7 Conclusion

An accurate, low-burden means of measuring and monitoring quality-adjusted outputs of publicly-funded social care services would provide a valuable tool that would allow a variety of stakeholders to evaluate productivity and value for money of social care interventions and how these are changing over time. It is essential that any such measure should reflect as far as possible the main objectives of social care: improved well-being and quality of life for service users. This study was testing the application of a general framework for establishing quality adjusted outputs or the value of care provided to one of the most resource intensive groups: residents of care homes. Ideally, in linking this to information that could be made available as part of the regulatory process, burden would be minimised and the value of information collected increased. The innovative approach to measuring outcome was applied across different and challenging service user groups. In the next chapter we describe the methods we developed and employed.

2. Method

The overall aim in developing the methodology was to use sound, valid and reliable techniques to establish resident outcomes and other key factors, such as quality of the care homes, that could be legitimately linked to independent measures of resident dependency and quality that either were or could be collected and generated by the regulator. Our key hypothesis was that the quality indicators provided by the regulator ought, particularly with the shift to a more outcome-based approach to regulation, to be associated with resident outcomes.

Previous approaches to identifying outcome and capacity for benefit (CfB) have been based on interviews with mentally intact (usually older) people in their own homes. In this situation people are able to reflect on what their levels of need were likely to be in the absence of services (Netten et al., 2006). As we identify in Chapter 1, a key challenge of this project was exploring whether it is possible to extend the measurement of outcomes to different settings and service user groups. The methods reflected the implications of focusing on people in a residential setting with a very different level and complexity of impairment, including learning disabilities and dementia.

We start by describing the overall research design and the approach taken to sampling the homes and residents. We outline the fieldwork process and identify the measures used before describing the testing of the methodology developed and the approach to identifying outcomes. The nature of the research was such that ethical considerations were of particular relevance. Appendix A provides information about the processes adopted for obtaining consent, ethical review and governance approval.

2.1 Research design and sampling

The aim was to achieve a final sample of care homes that reflected the full range of quality and types of home for older adults (OA) and people with learning difficulties (LD) across England. The objective was to conduct the fieldwork as contemporaneously with the inspection on which quality ratings were based, as was practical and manageable by the homes. Ideally we hoped to achieve 200 homes (100 for OA and 100 for people with LD) across four regions, reflecting a range of areas in terms of urban/rural and socio-demographic characteristics

(North West, West Midlands, London and South East). Homes were randomly selected from those inspected/due for inspection between January and December 2008 (see Towers et al., 2008 for details of the process). Care homes rated as excellent in their previous inspection were not officially due an inspection during this period so an alternative approach was needed to ensure an adequate sample of these homes. Additional data were supplied about these homes by the regulator and a sample randomly selected with the intention of achieving a sample of 20 such homes in total.

The observational nature of the method meant that there was a limit on the number of residents that could be included in the ‘focal’ sample, for which detailed information was collected. All homes are required to return an Annual Quality Assurance Assessment (AQAA) form, which, in addition to self-ratings of quality and other descriptive information, includes some basic (anonymised) information about up to 25 residents of the home, although it was very unusual to see as many as 25 residents listed here, even in larger homes. The exact number of residents was determined by the guidance given to the home by the regulator, and we were not privy to that information. A maximum of five residents were selected in each home from this form. As one of the aims of the study was to explore how well the needs of people from minority cultural groups are met in care homes, we wanted to ensure that we included in our sample residents from ethnic groups other than white British. These groups are very much a minority in care homes so we sought to achieve this by, whenever one or more of the residents was from a black or minority ethnic (BME) group, starting the random selection process from a BME resident. Beyond this there were no other selection criteria.

In addition to the in-depth information collected about focal residents (see below), basic information was collected about all other people identified on the AQAA who were resident in the home. In homes for people with LD, residents listed in the AQAA often constituted the entire home population, with rare exceptions such as larger homes or when someone had moved into the home since the most recent AQAA had been sent to the regulator. In homes for older people, which are generally larger, this would often constitute a sample of residents. There was little specific guidance provided to home managers on selection of residents for inclusion on the form; moreover, sometimes the inspector might request certain residents or people with certain characteristics should be included, so we are less confident that these are representative of all residents in the larger homes. In Chapter 3 we examine the information available about the wider population of residents in the homes and compare that with our sample.

Information about the inspectors who conducted the inspection and what they used in the course of each home's inspection was collected from inspectors in three waves after the majority of the fieldwork had been completed. After the fieldwork the regulator sent a data-set with basic information about all homes in England including dates of inspections and quality ratings. Where there were missing or dated data this was supplemented with information from the website.

2.2 Fieldwork process and data collected

Figure A.1 in Appendix A shows the process adopted for contacting and recruiting the homes. Thirteen fieldworkers conducted the fieldwork which took place over the course of two full days in each service, during which time the majority of information was collected about the home and residents. Data were collected about and/or from residents, care homes, care home staff and inspectors using self-completion questionnaires, standardised interviews, observations and administrative data.

2.2.1 User characteristic questionnaires

User characteristic questionnaires were sent in advance of the fieldwork visit and completed by staff who knew the residents well. They collected detailed information about the focal residents' age, ethnicity and impairments, their health (physical and cognitive), mobility, and care and nursing needs. The measures and data collected included:

- The short form of the Adaptive Behavior Scale Part 1 (Hatton et al., 2001; Nihira et al., 1993), the Quality of Social Impairment question from the Schedule of Handicaps Behaviours and Skill (Wing and Gould, 1978) and the Aberrant Behaviour Checklist (ABC) (Aman et al., 1995) which were used to collect information on the level of skills and ability and presence of challenging behaviour.

- Problems in activities of daily living (ADLs), continence, mobility and instrumental activities of daily living (IADLs) that have been used in a number of studies (Darton et al., 1997; Darton et al., 1998; Wittenberg et al., 2000) and that could be used to generate a number of measures including simple counts of need, Barthel dependency measure (Mahoney and Barthel 1965; Rodgers et al., 1993)and the Katz index (Katz et al., 1970).
- The Minimum Data Set Cognitive Performance Scale (MDS CPS) (Morris et al., 1994), which was developed for older people with cognitive impairment and validated against the mini-mental state examination (Morris et al., 1994)).
- Mental state items on anxiety and depression used in previous research (Darton, 1986; Darton et al., 1998).
- List of nursing care needs used in previous research (Darton et al., 1997).
- Items on experience and identification of pain.
- Use of community health care and hospital services.
- Existence of a person-centred plan.
- Weekly fees, source of funding, and whether the individual was placed out of their original home area.
- The Index of Participation in Daily Life (IPDL) (Raynes et al., 1994), the Index of Community Involvement (Raynes et al., 1989) and the Choice Making Scale (Conroy and Feinstein, 1986)) were used to measure the involvement people had in everyday activities at home and in the community.
- A set of items on minority cultural needs.

With the exception of the Short Adaptive Behaviour Scale (SABS), the social impairment question, the full challenging behaviour scale (ABC) and the questions about out-of-area placement, the same questions were asked for both older adults and people with learning disabilities. These four questions were completed only for people with learning disabilities. Some of the items of the ABC were asked for older adults to ascertain whether staff rated people as showing behaviours which they found difficult to deal with, but the full scale was not applied.

In addition, brief, anonymous questionnaires were used to collect a subset of these data for all other residents listed on the AQAA form (see above). The items covered age, ethnicity, ADLs, MDS CPS, mental state, nursing needs, source of funding and fees.

2.2.2 Resident and staff interviews

Structured interviews were conducted with both staff and residents (where they were able to participate) to collect detailed information about our SCRQOL domains for the focal residents. Staff interviews covered the domains and their perceptions of residents' current SCRQOL and expected level in the absence of the care and support provided by the home in each of our domains. Resident interviews focused on residents' experiences of living in the home, their views about staff and the help they receive, and their current care needs. This last included the health-related quality of life measure EQ5D (EurQol, 2009) and self-rated health (Bowling, 1995).

Observational measures

The observational measures formed the core of the evaluation of capacity to benefit, quality and outcome for the focal residents and were put together in a toolkit, which had seven sections:

1. A two-hour observation of engagement in meaningful activities and relationships (EMACR) (Mansell and Beadle-Brown, 2005). This observation used momentary time sampling approach with a 1 minute interval and rotation around focal residents every five minutes. The observation was usually conducted between 4pm and 6pm on the first afternoon, as this period often offers many opportunities for staff to interact and support the residents and maximised the possibility of seeing engagement and being able to rate the quality of staff support. However, research has shown that in terms of engagement, this time period is representative of the rest of the day – if people are engaged during this period they are generally also relatively well engaged at other points in the day and conversely if people were not engaged at this time then they were unlikely to be engaged at other periods (Mansell and Beadle-Brown, submitted). .
2. The Active Support Measure (ASM) and observation of support for choice making and communication (Mansell and Elliot; 1996, revised Mansell, Elliot and Beadle-Brown, 2005). These measures focus on the quality of the support staff provided to people around participation in meaningful activities and interactions and were completed for each service user at the end of the two-hour observation to complete the EMACR measure.

3. The Homelikeness Scale (taken from the Revised Residential Services Setting Questionnaire, the Welsh Centre for Learning Disabilies, 2003) was used to rate how homelike the home was in terms of the bedrooms, living and dining rooms, bathrooms and garden. In addition, a measure of the quality of the physical environment was completed at service level during the course of the fieldwork to provide information about the quality of the physical home environment (Sloane et al., 2002).
4. An adaptation of the Observable indicators of *Nursing Home Care Quality Instrument* (Aud et al., 2006; University of Missouri MDS and Quality Research Team, 2005). This was mostly collected at service level, but some questions were added to collect information for each of the service users. This scale collects multidimensional information about care quality, including interactions between staff and residents, type of care observed and even ratings of aspects of the physical environment not picked up in section three. Throughout the report we have used the term Nursing Home Quality Scale (NHQS) for our adapted scale.
5. Health: Pain assessment measure based on *Checklist of Non-verbal Pain Indicators (CNPI)* (Feldt, 2000); observations about diet/nutrition and exercise/activity.
6. A set of items reflecting minority cultural and ethnic needs. This was completed only for focal residents from BME groups.
7. Ratings of current SCRQOL and expected SCRQOL in each of our domains (see Chapter 1) in the absence of the care and support provided by the home. Fieldworkers were asked to complete this at the end of or immediately after the two-day fieldwork period.

Inspector and inspection data

Inspectors from CSCI were asked to complete a very short questionnaire detailing their own professional and academic background and what evidence they used to make their judgement about the home’s quality rating. Inspectors sometimes use information not publicly available when making their final assessments of the quality of the home. For example, they may be aware of complaints against the home or there may be adult protection issues that have influenced their rating. Inspectors also identified whether they had been trained in the use of Short Observation Framework for Inspectors (SOFI), and whether they had used it in this particular inspection.

Information about each home's performance in terms of national minimum standards, outcome or KLORA group ratings and overall star ratings were provided by the regulator, either directly or through the information available on their website. In addition to rating each home in terms of whether the national minimum standard has been met (on a four-point scale of exceeded, met, nearly met and not met), inspectors now rate each home in each of the domains of the standards on an additional scale of excellent, good, adequate and poor.[4] There are eight domains for homes for younger adults and seven domains for homes for older adults:

- Choice of home
- Individual needs and choices (Younger adults (YA) only)
- Lifestyle
- Personal care and health care support (YA)/ Health and personal care (OA)
- Concerns, complaints and protection
- Environment
- Staffing
- Conduct and management of the home (YA)/ management and administration (OA).

The ratings used in the research were based on the inspection that was most contemporaneous with the fieldwork.

2.3 Training and testing the process

The intensity of the data collection process and reliance on fieldworker judgements for our key outcome measures meant that piloting, training and testing for inter-rater reliability were of key importance.

Prior to the main stage of fieldwork a pre-pilot exercise was conducted with two homes (one OA and one LD) already known to the research team, to test the early drafts of the measures and observational processes in care home settings. The aim of this pre-pilot was to aid the development of the measures for the pilot study. An important element of this was to identify and make use of service

---

[4] http://www.cqc.org.uk/guidanceforprofessionals/socialcare/careproviders/inspection/keylinesofregulatoryassessment.cfm

users' comments and observations of the process and toolkit to improve the measure for the pilot stage.

For the pilot, six care homes were recruited opportunistically through the two regions that had already agreed to take part in the main study (North West and West Midlands). The full process was tested, including inter-rater reliability and feedback from homes and inspectors invited and taken on board in finalising the data collection instruments and fieldwork processes (Towers., 2008).

Intense training over two days was provided for all fieldworkers in March 2008, covering all aspects of the methodology and instrumentation from the importance of gaining consent through to interviewing skills and observational methodologies. After this initial training, fieldworkers were provided with a resource pack (training manual) and were paired with one another as part of a 'buddy' system. Each fieldworker was accompanied by their buddy in the first care home they visited, so each fieldworker visited two homes paired with another person before they began doing fieldwork on their own. A 'refresher' training day was held in June 2008, after all fieldworkers had visited at least one care home, to deal with issues that had arisen. Feedback was also given to researchers following preliminary analysis of inter-observer reliability.

2.4 Measuring SCRQOL and outcome

The data collection was intense in order to ensure that fieldworkers had as much information as possible on which to base their ratings of residents' experienced SCRQOL and expected SCRQOL in the absence of the care and support provided. For each of the eight domains residents were rated as:

- No need
    The individual has no or the type of temporary trivial needs that would be expected in this area of life of someone with no impairments
- All needs met
    The service and any other support provided ensures that the individual has no or the type of temporary trivial needs that would be expected in this area of life of someone with no impairments
- Low needs

Needs are sufficiently important or frequent to affect an individual's quality of life but do not have mental or physical health implications

- High needs

  Needs that would have mental or health implications if they were not met over a period of time

Appendix B shows how these general definitions were applied to each of the domains for both experienced and expected SCRQOL. These definitions drew both on the way in which people were asked about their self-ratings in interview and self-completion versions of the instrument, and on the nature and type of help that fieldworkers observed being provided. For the reasons we gave in Chapter 1, there was no 'expected needs' rating for the Dignity domain, which at this stage was also at a earlier stage of development in terms of definition than the other domains – relating more to the way that people were treated than how this appeared to make them feel (see section 4.5).

Fieldworkers were asked to provide evidence for their ratings for each domain for each resident. This was in part to provide a validation check that fieldworkers were applying the principles and guidance consistently. It was also in order that we could narrow down what evidence was needed for making judgements so in any future applications the fieldwork could be less intensive.

In scoring the SCRQOL measures, we wanted to reflect the relative importance of the different domains and levels of need reflected in composite SCRQOL states. Preference weights (see Appendix C) were drawn from a pilot preference study conducted for another research project (Burge et al., 2006).

2.5 Analyses

A variety of types of analysis were conducted covering the description of the samples, representativeness, reliability and validity of our measures, comparison of the OA and LD samples, and multivariate analyses, exploring the relationships between residents' capacity to benefit and outcomes and individual resident and home characteristics.

The descriptive analysis in Chapter 3 provides background information about the key characteristics of the homes and residents in the sample. For the homes that includes star ratings, standards met, ratings on the eight KLORA domains and home type and specialism. The descriptive information was also used to compare this sample to other samples of homes and residents, in particular to comment on the representativeness of the sample used in this study. Differences between the sample of older adults and the sample of people with learning disabilities in terms of their characteristics and the quality of their experiences and the ratings of outcomes were explored using a mixture of chi-square, independent T-tests and Mann-Whitney U tests depending on the level of measurement of data and whether the data in each analysis met parametric assumptions.

Statistical analysis of the representativeness of the sample homes compared with all homes inspected by October 2008 was conducted using chi-square analysis. In addition, characteristics of the focal sample were compared to a larger sample of residents from the same homes but not included in the detailed data collection, using a mixture of chi-square, independent T-tests and Mann-Whitney U tests depending on the level of measurement of data and whether the data in each analysis met parametric assumptions.

Reliability and validity analyses, described in Chapter 4, focused primarily on our measures of SCRQOL, using Kappa and percentage agreement to investigate inter-rater reliability and Cronbach's Alpha to identify the internal reliability of the measure. Correlations, including Spearman's correlation coefficient, were used to evaluate validity. Qualitative analysis, which was also used to investigate reliability and validity, is also described in the chapter.

A series of multi-level modelling regression analyses was used to explore the factors that contributed to capacity to benefit, outcome, and the relationships between outcome measures and inspection ratings. Due to the complexity and variety of analyses conducted, specific information on the models and tests used are described in detail in Chapters 5 and 7.

3. The sample

Rather than obtaining a nationally representative sample, the key objective was to ensure we included a range of home types and quality, resident characteristics and dependency, and had an understanding of the characteristics of inspectors and inspection processes that underpinned the regulator quality ratings. The intensiveness and nature of the fieldwork (described in Chapter 2 and Appendix A) was such that we anticipated a relatively low response rate at the home level. In addition, responses at the resident level might be expected to be affected by problems in obtaining consent and permission, particularly for those who lacked mental capacity, and by the vulnerable nature of residents, many of whom would be very ill. While the overall objective was not necessarily to be nationally representative, clearly it is important that we understand the nature of the samples of homes and residents and any potential sources of bias.

Here we describe the response rates and the characteristics of the final samples of homes and residents, inspectors who conducted the inspections, and the inspection processes used. Where possible we compare the information about our sample with other sources to identify how representative it is of the national picture, or at least to put our sample in the wider context.

3.1 Homes response and characteristics

The aim was to achieve a final sample of 200 care homes (100 older adults and 100 homes for people with intellectual disabilities) across four CSCI regions (North West, West Midlands, London and South East). Homes were randomly selected from those inspected/due for inspection between January and December 2008 (for a summary of the sampling process please see Appendix G; a more detailed account is available in Towers et al. 2008). Care homes rated as *excellent* in their previous inspection were not officially due an inspection during this period so an alternative approach was needed to ensure an adequate sample of these homes. Additional data were supplied about these homes by CSCI and a sample randomly selected with the intention of achieving a sample of 20 such homes in total.

From a sample of 444 homes, we achieved co-operation from 173 homes: 83 homes registered for older adults (OA) and 90 homes registered for people with a learning disability (LD). The overall homes response rate was 39 per cent, with response from LD homes a little higher than for OA homes.

Table 3.1 Homes sample and response

	All homes	Older Adults (OA)	Learning disability (LD)
Selected sample	444	235	209
Consent to participate not obtained	270	151	119
Later dropped out	1	1	0
Responding homes	173	83	90
Response rate (%)	39	35	43

The homes were spread across the four regions, with the largest proportion, 35 per cent (60 homes), in the South East, nearly a quarter (42 homes) in London and the remainder equally divided between the West Midlands (36 homes) and the North West (35 homes). For the most part homes were fairly evenly divided between OA and LD, although there were more homes for OA in the North West (28) and for LD in London (29). This partly reflected the distribution of homes that the regulator identified as being inspected during the period (more inspected OA homes in the North West) and partly differential response rates (57 per cent of London LD homes that were approached participated).

3.2 Examination of possible non-response bias

While not unexpected given the demands of the research, the homes response rates were lower than we would have liked. To investigate possible non-response bias, data provided by the regulator were analysed to compare homes participating in the study with those which were approached but for which consent to participate was not obtained. The analysis is based on the 174 homes initially recruited. To further investigate the composition of the sample, regulator data were analysed to compare sampled homes with those not selected. Data were available on overall quality ratings, percentage of standards met, and type and size of home in England as a whole and for our four regions. We also draw on data reported in the State of Social Care 2009 report published by the regulator (Commission for Social Care Inspection, 2009). Table 3.2 summarises these comparisons for all homes, OA and LD.

Table 3.2: Comparison of care homes approached and not approached, participating and non-participating

	Participating homes (n=174)	Approached but did not participate (n=269)	Total sampled 4 regions only (n=418)	Not sampled 4 regions only (n=8869)	Total 4 regions (n=9285)	Total all regions (n=17676)	National from State of Social Care report[1]
Number of places							
Mean	21	23	22	24	24	25	NOT PROVIDED
(range)	(2-114)	(1-150)	(1 – 150)	(1 – 253)	(1-253)	(1-253)	
Quality							
% No stars	2	4	3	2	2	2	3
% 1 star	22	28	26	16	16	17	24
% 2 stars	51	40	44	59	59	58	56
% 3 stars	14	9	11	15	15	15	13
% no rating[2]	11	19	16	7	8	8	4
% standards met	82	78	79	86	84	84	NOT PROVIDED
% standards exceeded	5	4	5	8	7	7	NOT PROVIDED

1. Commission for Social Care Inspection (2009). Covered 2008/09 the period of the data collection; 2. Not rated, rating suspended, data missing.

Table 3.2 illustrates that overall our sample was closely matched to the national picture presented in the State of Social Care Report 2009. Star ratings were available for the whole sample bar one home for people with LD that had not been rated by the end of the fieldwork period. Despite repeated chasing, full information was not available on the standards for 11 per cent of the sample. Excluding the non-rated homes, sixteen per cent were rated ‘excellent’, compared with 14 per cent nationally.

As was expected, there was a slight tendency for non-responding homes to be more likely to be rated as poor or adequate, 32% compared with 24% of the responding homes. With almost a quarter of the responding homes in the poor/adequate category, however, we have a reasonable representation of lower quality rated homes in the sample. Chi-square analysis showed no significant differences in quality ratings between the participating homes and those who were approached but refused to take part. This was also true when participating homes were compared with homes not sampled in the four regions included in the study (North West, West Midlands, South East and London).

The homes in the sample met 82 per cent of standards (ranges 27-100) overall. On average, homes only exceeded 5 per cent of standards (0 – 63). No home exceeded all standards. The percentage of lifestyle standards met was slightly lower: 69 per cent (0 – 100). On average, just 7 per cent of lifestyle standards were exceeded. There were no statistically significant differences between the sample of participating homes and the homes that were not approached in terms of the percentage of standards met and the percentage of standards exceeded. However, homes who participated exceeded significantly more lifestyle standards (7 per cent) compared with those approached but not participating (4 per cent) ($z=3.095$, $p<0.01$, $n=437$) which reflects the fact that poorer homes were less likely to consent. Those not approached met more standards than participating homes ($z=2.895$ $p<0.01$) but the actual difference in terms of number of standards was very small.

The comparisons between responding and non-responding homes are provided separately for the LD and OA homes in Table 3.3. With regard to quality ratings, the same patterns of non-response were found for each of the two types of home as for the whole sample but none of the differences between the participating homes and those not sampled or those sampled but not participating was statistically significant. None of the learning disability homes included was rated as ‘no stars’. We had attempted to recruit twice as many of these services as would be expected from the whole sample, allowing for a 50 per cent agreement rate, but all of the LD homes rated as poor refused to take part.

Tables 3.2 and 3.3 show there was no significant difference between the size of those services which were selected and participated compared to those selected but for which consent was not obtained, either for the sample as a whole or for LD or OA homes separately

Table 3.3: Comparisons of sampled, non-sampled and participating OA and LD care homes

Characteristic of home	**Not selected**	**Approached but did not participate**	**Participating homes**	**State of Social Care report[2]**
LD homes	Mean = 9	Mean = 9	Mean = 7	All LD homes
Number of places	(1 to 252)	(1 to 73)	(2 to 21)	Mean = 9.6
OA homes	Mean = 35	Mean = 35	Mean = 35	All OA homes
Number of places	(1 to 253)	(3 to 150)	(3 to 114)	Mean = 34.7
Quality rating LD services				
N	5386	100	78	-
% No stars	3	3	0	2
% 1 star	18	36	26	22
% 2 stars	63	48	59	61
% 3 stars	16	13	15	15
% standards met	86	79	82	82
Quality rating OA services				
N	9288	121	73	-
% No stars	3	5	3	4
% 1 star	18	35	25	28
% 2 stars	63	49	57	55
% 3 stars	16	12	16	12
% standards met	83	77	82	85

1. For consistency with other information in the table, the number of places taken from data provided by the regulator.

2. For the State of Social Care Report figures, 1.1 per cent of older adult services were not yet rated; 6.1 per cent of younger adult services were not yet rated.

There was no difference between LD and OA homes in terms of the percentage of standards met overall. OA homes were rated to exceed slightly more standards (average 6.46) than LD homes (3.92). However, LD homes were more likely to meet lifestyle standards. On average, OA homes met 45 per cent of lifestyle standards compared to 92 per cent in LD homes. LD homes exceeded 8.48 per cent of lifestyle standards compared to 5.61 per cent in OA homes.

Table 3.4 presents the percentage of homes being rated as poor, adequate, good and excellent on each domain of the KLORA or NMS outcome groups. Only a small number of homes for older adults were rated on ‘Individual needs and choices’. As can be seen the majority of homes were rated as good on all of the domains, with very few homes rated as poor on any domain. There were significant differences (on a Mann-Whitney U test) between older adults and younger adults homes in terms of the ratings on three domains: Choice of home ($z=2.456$ $p=0.014$), Environment ($z=6.694$ $p<0.001$) and Staffing ($z=2.899$, $p<0.01$). In all of these, OA homes were rated more highly than LD homes.

Table 3.4: Percentage of sample rated within each category for each of the 8 KLORA domains

		Choice of home	Individual needs and choices	Life-style	Personal and healthcare support	Concerns complaints and protection	Environment	Staffing	Conduct management of the home
Overall sample									
n		167	97	167	167	167	167	167	167
%	poor	0.6	2.1	0.6	3.6	1.2	1.2	2.4	3
%	adequate	12	14.4	8.4	15	12	19.8	24.6	18
%	good	79	68	67.7	68.9	79	68.9	64.7	64.7
%	excellent	8.4	15.5	23.4	12.6	7.8	10.2	8.4	14.4
Homes for older adults									
n		80	11	80	80	80	80	80	80
%	poor	0	0	1.3	5	2.5	2.5	1.3	3.8
%	adequate	8.8	9.1	8.8	18.8	12.5	11.3	21.3	16.3
%	good	82.5	72.7	66.3	58.8	73.8	70	67.5	63.8
%	excellent	8.8	18.2	23.8	17.5	11.3	16.3	10	16.3
Learning disability homes									
n		87	86	87	87	87	87	87	87
%	poor	1.1	2.3	0	2.3	0	0	3.4	2.3
%	adequate	14.9	15.1	8	11.5	11.5	27.6	27.6	19.5
%	good	75.9	67.4	69	78.2	83.9	67.8	62.1	65.5
%	excellent	8	15.1	23	8	4.6	4.6	6.9	12.6

Size

Tables 3.2 and 3.3 show there was no significant difference between the size of those services which were selected and participated compared to those selected but for which consent was not obtained, for the sample as a whole or for LD or OA homes separately. However, those services which chose to participate were slightly but significantly smaller in size than those not selected ($t= 2.360$; $p<0.05$; $df = 18214$). This difference was only an average of four places and did not hold up when the analysis was repeated for just the four regions included in this study.

As we would expect, homes for OA were larger than those for people with LD. The mean number of places we recorded (which did not exactly match the data provided by the regulator for those homes) across the whole sample was 21 (range 2–114; SD 18.57). Some of the larger services had some places unfilled. The range of occupied places was 2–112 (mean 18). In OA homes the average number of places was 33 (range 3-114; SD 18.78). This was slightly smaller than the average nationally,[5] which was 35 (range 4-140). This difference was not statistically significant and probably reflects the omission of the very largest homes. Three homes in our sample (2 per cent) had 100 places or more; 7 (4 per cent) of the homes in our sample had more than 50 places. This is comparable with the national picture, taken from the dataset provided by the regulator in December 2008, in which 9 per cent of homes had more than 50 places and just under 1 per cent of homes had more than 100 places. Services for people with LD were substantially smaller, with a mean of eight places (range 2-21, SD 4.13).

Sector and specialisation

Table 3.5 summarises the data on the providing sector of the participating services. The majority of care homes nationally and in our sample are in the private for-profit sector. This is particularly true of care homes for older people. Nevertheless, a substantial proportion - nearly a fifth of OA and over a third of LD services - were provided by the third sector. This is slightly higher than for the national picture, which is not available by type of home and thus possibly skewed by the dominance of OA homes. Nationally 18 per cent of all care homes are provided by the third sector, with 74 per cent provided by private organisations (calculated from Table C2 in the State of Social Care Report 2009).

---

[5] Taken from data provided by the regulator.

Table 3.5: Proportion of care homes drawn from each sector

Sector	Whole sample (%) n=173	Older adults services (%) n=83	Learning disability services (%) n=90
Private	64	74	55
Voluntary	28	19	36
Local authority	5	6	3.4
NHS	1	0	2.2
Other	2	1	4.4

Just under a fifth (18 per cent) of all homes were registered nursing homes, but only one of these was for people with learning disabilities (this was for people with dementia). This is slightly lower than the national average of 22 per cent (calculated from Table C1 in the State of Social Care Report, 2009). In terms of the sample of focal residents, over a third (38 per cent) of the OA sample lived in a nursing home compared with just four people (1 per cent) with learning disabilities.

An analysis of the information available on the regulator website identified that a fifth of the sample homes were classified as ‘specialist’: that is, that they catered for particular ethnic, religious groups or for particular conditions (for example, dementia). Table 3.6 shows the distribution of homes across the different types of specialisms.

Table 3.6: Home specialisms

Type of specialism	Number	Percentage
No specialism documented	137	79
Religious specialism	13	8
Christian	10	6
Jewish	3	2
Condition	18	10
Dementia	7	4
Dual diagnosis[1]	5	3
Epilepsy	2	1.2
Sensory impairments	2	1.2
Profound and multiple	2	1.2
Other specialism	4	3

1. Dual diagnosis usually refers to homes that specialise for people with both learning and mental health needs.

There was an association between specialism and sector. For older adult services, 19 homes were classified as specialist. Just over half (53 per cent) were private, 37 per cent voluntary and 10 per cent LA. Among non-specialist homes, 79 per cent were private, 14 per cent were voluntary, 5 per cent LA and 2 per cent 'other'.[6] For LD services, 16 homes were classified as specialist, of which three quarters were voluntary, three (19 per cent) private, and one (6 per cent) 'other'. Non-specialist services were 63 per cent private, 20 per cent voluntary, 4 per cent LA, 3 per cent NHS and 3 per cent other. This pattern was statistically significant for LD homes (p=0.001) but not for older adult homes.

3.3 Characteristics of focal residents and non-focal

As we describe in Chapter 2, our focal residents were sampled from those residents identified on the AQAA form provided by the home. Table 3.7 shows the number of residents in the homes, identified on the AQAA forms from the homes, and included in our focal and non-focal samples. As we would expect, given the much smaller size of homes for people with LD, our data covered a much higher proportion of the resident population and there was a much lower number of 'non-focal' residents.

---

[6] Data supplied by CSCI – no definition of 'other' available.

Table 3.7: Resident sample

	Overall sample	Older adults services	Learning disability services
Number of residents in participating homes[1]	2990	2407	583
Number of residents identified on the AQAA forms	1729	1194	535
Focal resident sample (n)	741	375	366
Non focal resident sample (n)	810	675	135
Total sample (n)	1551	1050	501
% of all home residents in total sample	52	44	86
% of all home residents in focal sample	25	16	63

1. Total number of places reported occupied.

The focal resident response rate is not straightforward to estimate because when residents who had been selected were no longer resident or not able to participate for other reasons, another resident was selected to replace them. In total 860 people were approached for consent (via the home manager), from which the final sample of 741 people (86 per cent) participated. A higher proportion of residents with LD approached took part (94 per cent) than OA (80 per cent).

Table 3.8 shows the reasons for non-participation in the study at any stage of those selected for the focal sample. As we would expect, more residents of homes for OA had died or left the home than for homes for LD. Actual refusal rates were very low: initial refusals are included in ‘other’ reasons for non-participation. There was a similar pattern of reasons for lack of information about residents in the non-focal sample. Of the 59 for whom we had reasons, about half (48 per cent) had died and over a third (37 per cent) had left the home.

Table 3.8: Reasons for non-participation in focal sample

	Overall sample (n)	Older adults' services (n)	Learning disability services (n)
Died	40	39	1
Left home	12	6	6
In hospital	20	18	2
Other	10	3	7
Refused on the day of the visit	20	10	10
No information	17	17	0
Total	119	93	26

Table 3.9 shows the characteristics of the overall sample by whether or not the residents were in the focal sample and service user group. There were few statistically significant differences between the focal and non-focal sample. In both OA and LD homes, the focal residents were slightly younger ($p<0.05$ and $p<0.01$ respectively). OA focal residents also had slightly higher Minimum Data Set Cognitive Performance Scale (MDS CPS) scores ($p<0.01$) suggesting that the most impaired were less likely to participate. However, the full range of cognitive functional ability as measured by the MDS CPS was represented in our sample.[7]

---

[7] Excluding a 'coma' rating that is available on the full scale.

Table 3.9: Characteristics of resident sample

	All services		Older adults services		Learning disability services	
	Focal	Non-focal	Focal	Non-focal	Focal	Non-focal
Sample size	741	810	366	677	375	132
Age	65	80	84	86	46	49
(mean and range)	(19-104)	(21-104)	(31–104)	(37-104)	(19-100)	(21-84)
% male	43	32	30	27	58	57
% white British	87	95	91	96	84	89
ADL 9 item scale[1]	8.47	9.86	10.01	10.43	6.87	6.04
(mean and range)	(0-18)	(0-18)	(0-18)	(0-18)	(0-18)	(0-18)
MDS CPS[2]	2.95	2.53	2.58	2.46	3.37	2.53
(mean and range)	(0-6)	(0-6)	(0-6)	(0-6)	(0-6)	(0-6)

1. High score = high need. Each item scored: 0 no problem, 1 with difficulty, 2 only with help. Scale includes: washing, bathing, dressing, use WC, get in and out of chair/bed, move in bed, use WC, feed self, mobility indoors.

2. High score = more cognitively intact.

Table 3.10 shows additional information that we had about our focal sample, including sensory impairment, diagnoses of specific conditions, and the widely-used health-related quality of life (HRQOL) measure, EQ5D. In addition to the information shown, 4 per cent of the OA sample was recorded by staff as having a learning disability.

Table 3.10: Focal sample resident characteristics

	Whole sample (n=741)	OA homes (n=375)	LD homes (n=366)
% with physical disability	35	38	34
% with hearing impairment	18	26	10
% with a visual impairment	24	31	16
% with mental health needs	32	40	23
% with autism diagnosis	10	0.3	20
% with epilepsy	17	3.5	31
EQ5D[1]	0.47	0.42	0.53
(Mean and range)	(-0.43 – 1)	(-0.43 – 1)	(-0.36 – 1)

1. EQ5D is a preference-weighted measure with 0 anchored to ‘dead’ and 1 to ‘good health’. Negative scores are associated with health states worse than dead.

An important aspect of this study was using the same measures as far as possible for OA and people with LD. In the tables and in the analyses in subsequent chapters, we have focused on the use of the ADL, EQ5D and MDS CPS measures (see Appendix D for the distributions of the individual items of the scales). ADLs and MDS CPS are most frequently used to measure abilities of OA. For people with LD, measures such as SABS, or indeed a full measure of Adaptive Behaviour Scales, are more frequently used. We found that ADLs were highly-correlated with SABS, particularly a 12-item ADL measure, which, in addition to the nine items shown in Table 3.9, included ability to use a vacuum cleaner, wash clothes by hand and manage personal affairs (r=-0.84).[8]

Homes provide the regulator with information about the number of their residents who have certain needs or characteristics on their AQAA form. The initial aim was to use this as a basis for comparing with the characteristics of all residents in the sample; however, the reported figures gave rise to some concern. Only 75 per cent of residents in the LD homes were identified as having a learning disability. A much higher proportion of people in our sample were identified as having mental health needs or a physical disability than in the AQAA (20 per cent (including dementia) and 17 per cent respectively). If anything, as we discuss below when comparing our sample in OA nursing homes, our concern might be that we have a more able rather than less able sample than other residents in the home. Among LD homes, our sample includes such a high proportion of residents, the difference would suggest under-reporting of residents' impairment to the regulator rather than that our sample is particularly highly impaired.

As the majority of our analyses were conducted on the focal sample our interest is primarily in how representative our focal sample is in terms of the wider care home resident population. This is difficult to establish because of lack of information about the wider resident population, but we can draw on previous studies to provide some context.

---

[8] The 12-item ADL was also correlated significantly with the same measures of residents' experiences that SABS was correlated with and which are reported in later chapters: engagement, active support, staff ratio, adapted communication and supporting choice, score on the Index of Participation in Daily Life (IPDL), the Index of Community Involvement (ICI) and the Choice Making Scale (CMS). Regression analysis exploring the predictors of engagement in which SABS is normally a predictor, found that the 12-item ADL also predicted level of engagement.

Older adults

For OA homes, information based on good sized representative data that does exist is dated and we know that in OA homes dependency has been increasing. The characteristics of publicly funded older adults moving in to care homes in 2005 were the equivalent to those moving in to nursing homes in 1995 (Netten et al., 2006). We draw on a national survey of over 11,000 residents conducted in 1996 (Netten et al., 2001). A survey of publicly funded admissions conducted in 2005 (Darton et al., 2006) also provides us with helpful insights. We also draw on a very recent but small scale study (n=59; 20 older people in homes for older people, 19 people with learning disabilities in homes for older people and 20 people with learning disabilities living in learning disability services) reported in Higgins and Mansell (2009) and the results of the Partnership for Older People Projects (POPP) evaluation which included a sample of 68 residents of care homes (Windle et al., 2009).

In terms of age, the OA sample was very similar to other surveys: 84 years compared with 85 in 1996 and 82 in the Higgins and Mansell study. The proportion of male residents (30 per cent) was slightly higher than in previous surveys but this might partly reflect an increasing proportion of male residents over time. The proportion was 22 per cent in 1996, 26.7 among admissions in 2005, and 35 per cent in the 2009 Higgins and Mansell sample. Historically data have not been good on ethnic origin of residents. However, at 9 per cent, our sample had the same proportion of OA from black or ethnic minority groups as were identified in the sample of publicly funded admissions in 2005 (Darton et al., 2006).

When we make our comparisons of dependency-related characteristics we need to distinguish between homes which are registered for nursing and general care homes as the populations are very different. Previous surveys have used Barthel (Rodgers et al., 1993) as a summary dependency measure. This ranges between 0 and 20, with low scores indicating higher levels of dependency. The average score for our sample in personal care OA homes was 12.5, compared with a mean of 14 in 1996, reflecting the more dependent population we would expect. Our nursing home sample was significantly more dependent than the personal care home sample, with an average score of 7.4 ($p<.001$). However, this was not much more dependent than the nursing home sample in 1995, which had an average Barthel score of 7.7 (Netten et al.,2001). The average Barthel score for publicly funded people admitted to personal care homes in 2005 was 10.4 and to nursing homes was 5.4, both lower than our home population samples. However, average EQ5D scores were virtually identical to those reported in the POPP evaluation in 2008 (0.51 in personal care homes and 0.28 in nursing homes compared with 0.50 in our sample in personal care homes and 0.27 in nursing homes).[9]

In terms of cognitive impairment there was a similar picture. Using the MDS CPS, 13 per cent of our sample in personal care homes were classified as intact (scoring 0) compared with nearly a third (31 per cent) in residential care in 1996. In terms of severe impairment (scoring 4-6), in our personal care home sample 43 per cent were severely impaired compared with 23 per cent in 1995. However, in nursing homes, 21 per cent of our sample was intact compared with 17 per cent in 1995. The proportion with severe impairment was the same, at 40 per cent. Among publicly funded people admitted in 2005, 15 per cent of those admitted to both nursing and personal care homes were intact, and 40 per cent and 54 per cent respectively were identified as severely impaired.

People with learning disability

In terms of age, our focal sample matched other LD samples well (46 years compared to between 43 and 47 years in other studies). Beadle-Brown and colleagues[10] found a mean of age of 47 years (range 22-90) in a sample of 120 individuals with learning disability in residential or supported living services provided by a national charity. In terms of gender, our sample was also comparable: 58 per cent were male compared with 54 per cent in the Beadle-Brown et al., (2009) sample; and 57 per cent in the Beadle-Brown et al., (2008b) sample of 29 individuals living in residential care provided by a charity in the South East of England. When compared to other LD samples, our sample included more people from an ethnic background other than white British. A higher proportion of residents in our sample were from BME groups: 84 per cent were identified as white British, compared with 96 per cent in the other studies of people with LD in residential care (Mansell et al.; 2008; Beadle-Brown et al.; 2008b). This is likely to be the effect of our sampling policy. As described in

---

[9] The difference in EQ5D between personal care homes and nursing homes was statistically significant in our sample ($p<.001$). In the general population average EQ5D score for those aged 65-74 is 0.78 and for those aged 75 or more is much lower, 0.73 (Kind et al 1999). There are no data available for people with learning disabilities.

[10] Analysis of unpublished data.

Chapter 2, we wanted to ensure a large enough group from ethnic minorities to facilitate analysis of needs of cultural minorities in care homes.

Our LD sample was also comparable to other samples in terms of the percentage with physical disability (34 per cent compared to 37 per cent found by Beadle-Brown and colleagues (2008b) and 35 per cent in the later study (Beadle-Brown et al., 2009)): with mental health needs (23 per cent compared to 24 per cent in Beadle-Brown et al.,(2008b) and 28 per cent in Beadle-Brown et al.,(2009)): with epilepsy (31 per cent compared with 33 per cent in both the comparison samples) and with autism (20 per cent compared with 17 per cent in Beadle-Brown et al., (2008b)). As with dementia and OA, it is important to note that it is difficult to ask about autism in adults with LD: not very many people will have a diagnosis even though they may show many of the features of autism. Emerson noted that autism is often under-reported by staff. If staff are asked to say how many people are socially impaired (an indicator of autism), they also find this very difficult to do and tend to report numbers between 60 and 80 per cent.

In terms of level of behaviour, the sample was also consistent with other studies. The mean Short Adaptive Behaviour score (SABS) score was 39 per cent, matching the mean score of 39 in the Beadle-Brown et al., (2009) study and is only slightly lower than the 46 per cent found in the smaller sample in Beadle-Brown et al., (2008b). The range (0.88-95.61 per cent) indicated a full range of level of ability: people whom staff judged to be completely dependent on staff for any of the basic personal care needs through to people who could do many things for themselves. The Aberrant Behaviour Checklist reflects challenging behaviour, with a lower percentage score reflecting higher levels of challenging behaviour. The average score was 12 per cent for residents with LD, ranging from zero to 47 per cent. This sample showed slightly more challenging behaviour than most other samples studies, where ABC percent score ranges from 17 in the Beadle-Brown (2008b) sample to 25 in the Mansell et al., (2008) study.

Fees and funding

We collected information about sources of funding and levels of weekly fees. Overall, 76 per cent of residents were publicly funded either wholly or in part. A significantly higher proportion of people with LD were publicly funded (86 per cent) than OA (55 per cent) ($p<0.001$). Excluding the fees reported for residents of local authority-run homes, as these reflected the resident contribution rather than the full cost, the mean level of fees for our full sample was £861, ranging from £297 to £3427. Compared with OA homes where the average weekly fee was £525 (ranging from £300 to £3384), fees were higher for people with LD: £1218 per week on average (range £297 to £3427). Fees for OA were significantly higher in nursing homes (£588) than general care homes (£469) ($p<0.001$). These compared with average weekly fees for OA homes nationally of £678 and £467 per week in 2008 (Curtis, 2008). Equivalent fee information is not available for homes for LD, but estimated weekly unit costs of group homes were estimated as £1,131 per resident (Curtis, 2008).

3.4 Characteristics of inspectors of homes in study

Information was collected about the background, experience and qualifications of inspectors. Questionnaires were returned from inspectors for 121 homes. The background of inspectors of the homes is summarised in Table 3.10.[11]

---

[11] There is a certain amount of double counting as the data are reported at the level of the home. In total there were 71 (LD) and 69 (OA) inspectors of homes in the sample. The majority of inspectors inspected just one home, but 16 learning disability inspectors inspected more than 1 home, and one inspector had reviewed 5 homes. Twelve inspectors of OA homes inspected more than one home in the sample, with a maximum of three homes inspected by the same inspector.

Table 3.10: Characteristics of inspectors of homes in the sample

Inspector Characteristics	Percentage
Length of time as inspector	
5 years or less	29
6-10 years	51
More than 10 years	20
Qualifications[1]	
Nursing	39
Social work	25
Mental health related qualification	12
Education	6
Management	16
Other (not including NVQ)	18
Highest level qualification	
Degree	26
Diploma/certificate	46
NVQ	3
Registered nurse	22
Other	3

1. Some inspectors had more than one qualification.

Nursing was the most frequently held type of qualification, but many had experience in social care. Nearly a third (30 per cent) of homes were inspected by inspectors who had been a residential home manager, 15 per cent a social services team manager and another 4 per cent had been an area manager within social services. A quarter of homes were inspected by people who had a health services background: 14 per cent had been a nurse and 11 per cent had been a ward manager or held another health based team leader position.

There was some association between educational and employment background and type of home inspected. Inspectors of OA homes were more likely to have a nursing background ($p<0.01$). In terms of previous jobs inspectors of OA homes were more likely to have been a nurse before becoming an inspector, whereas inspectors of LD homes were more likely to be a team or area manager in social services ($p=0.004$).

Inspectors had a lot of experience to draw on in inspecting homes. Nearly two thirds (61 per cent) of homes were inspected by someone who had been an inspector for five years or more and a fifth over ten years. Nearly all had previous experience of inspecting OA homes (99 per cent); of inspecting LD homes (93 per cent) and of inspecting nursing homes (88 per cent). There was no difference between those who had inspected OA and those who had inspected LD homes in our sample on whether they inspected older people's homes. However, inspectors of OA home in this sample were less likely to inspect LD homes in general ($p=0.003$). In addition, inspectors of LD homes in our sample were less likely to inspect nursing homes ($p=0.006$).

We were particularly interested in whether inspectors had training and experience in use of the observational tool that has been developed for inspectors to identify the well-being of residents, particularly those whom it is difficult to interview. Nearly all the inspectors had received (94 per cent) Short Observation Framework for Inspectors (SOFI) training, but 43 per cent had never used SOFI; 21 per cent had used it once, 9 per cent twice and 8 per cent three times. Just 17 per cent had used SOFI more than three times.

3.5 Inspection processes

Inspectors were asked what type of information they had gathered as part of the inspection and what they had drawn on in making the quality ratings. Table 3.11 shows all the types of information and the degree to which it was drawn on for OA and LD homes' ratings. Drawing on various records, talking to staff, residents and informal observation were frequently employed. The information and activities that they were most likely to say had an impact were sampling user records, talking to staff and informal observation of residents in the home, then case tracking, talking to residents and talking to managers. Inspectors of OA homes were more likely to say that a tour had a significant impact on their rating ($p=0.006$).

We had hoped to test whether the observational techniques inspectors used had an impact on the relationship between resident outcomes and SCRQOL and the regulator quality ratings. Unfortunately, the fact that most inspectors had been trained but very few made use of the technique in these inspections meant it was not possible to explore this.

Table 3.11: Inspection activities

Type of information	Used as part of inspection? % yes (n=122)	Significant impact on overall quality rating? % yes (n=93)
AQAA	97	50
Tour of premises	95	66
Talking to manager	89	80
Talking to staff	95	91
Talking to residents	91	85
SOFI	3	5
Informal observation of residents in the home	95	91
Experts by experience	3	4
Questionnaire to other professionals	46	24
Link resident	1	1
A communication toolkit	2	2
Records of complaints against the home	93	57
Notifications of significant events	98	59
Case tracking	92	87
A sample of service user records	94	92
Quality assurance monitoring	85	69

3.6 Conclusions

While clearly the response rate was lower than we had hoped, our sample is a good match to the national figures reported, in Table 3.2 and 3.3, for all younger adult and all older adult care homes. The only group where we have no representation is the very small group which has a zero rating for homes for people with learning disabilities.

In so far as it is possible to verify this, our sample of LD residents appeared to be typical of residents in our sample homes and in care homes nationally, with the possible exception of the most highly dependent. OA residents of personal care homes are, as we would expect, more impaired than in previous surveys, although the sample appeared more able than publicly funded admissions in 2005. There must be some reservations about the representativeness of our sample of nursing home residents, however. Although the comparison with the recent POPP study was encouraging, this was for a relatively small sample and we would have expected higher levels of dependency given the results of previous large-scale studies. The low rate of refusal of residents selected at random from the list of residents on the AQAA provided by the home suggest that the source of this bias is probably which residents were selected to be identified on the AQAA. We return to consideration of the implications of this in Chapter 8. Critically, however, our sample does include the full range of impairments to be expected in residents of OA and LD care homes.

There is no national information available on inspector or inspection characteristics, but our data provide an insight into the considerable level of experience of inspectors and what is used in practice when inspecting homes.

4. Reliability and validity

Cognitive impairment and communication difficulties are widespread among residents of care homes, especially homes for older adults (OA) and people with learning disabilities (LD). One of the aims of the project was to develop an approach to identify social care related quality of life (SCRQOL) and outcomes for care home residents that could be applied to all residents. The approach, as we described in Chapter 2, was based on fieldworkers gathering evidence from a variety of sources including: previously validated instruments, structured observation, and interviews with staff and, where possible, residents. On the basis of this evidence and guidance provided they were asked to make judgements about residents' current and expected SCRQOL in each domain. It is important to examine the reliability (that is, whether the judgements are consistent between observers and by observers over time, and whether the scale appears to reflect an internally consistent construct) and validity, (that is, whether the scale is associated with other measures in ways we would expect). This is both for the purposes of interpreting our results and for recommending the way forward in terms of the future.

We start by describing the methods that we used and go on to report on the analysis of the reliability of the individual domains and ASCOT measure of SCRQOL, both current and expected in the absence of services (Sections 4.2 – 4.4). We discuss particular concerns raised in the domain of Dignity and the evidence about rating OA and people with LD (Section 4.5). We report on the internal reliability of the ASCOT measures (Section 4.7) before briefly reporting on the reliability of other measures included in the toolkit (Section 4.8). While it is not feasible to separately test the validity of the ASCOT measure within this study, we can provide some insight through the association between these other measures and the ASCOT SCRQOL domains.

4.1 Overview of methods used

We employed both qualitative and quantitative approaches to evaluating reliability and validity. We examined inter-rater reliability (whether different raters make the same judgements under the same circumstances) and analysed individual fieldworker reliability in terms of the consistency of the basis for the SCRQOL ratings made. We examined the relative reliability of the different domains and whether there were differences in the approaches used and bases for rating SCRQOL for OA and people with LD (Sections 4.3 and 4.6). We also examined the internal reliability of the ASCOT measures of SCRQOL with and in the absence of services. For the other measures used in the study (for example standard measures such as the Active Support Measure and the Nursing Home Quality Scale) we first evaluated the inter-rater reliability on these measures (Section 4.7). We then examined the associations between these measures and ASCOT, as a way of assessing the validity of ASCOT as far as possible within the limitations of the data available (Section 4.8).

In order to test inter-rater reliability, two raters completed the entire toolkit for a subsample of the residents. All fieldworkers conducted at least four visits in pairs: two visits as the main observer and two as a buddy. On these visits, the fieldworkers had to complete all of the measures independently. They were allowed to discuss their ratings after they had completed their ratings but were not allowed to change them. They were, however, asked to make a note in the reflective record from the visit which highlighted anything they think they coded differently from their buddy. One set of buddy visits were conducted within the first two months of the study, the other set was collected in the latter stages of the study.

Inter-observer/rater reliability was available for 113 residents (the exact number of users included in each analysis varies slightly depending on the measure) in 28 services.

In the quantitative analysis, reliability was examined using Kappa Statistic and percentage agreement. Where appropriate Spearman's ranked order correlation coefficient(i.e. an ordinal scale) has also been calculated. The statistics have been calculated across all pairs of observers and then the mean agreement (and range) calculated for each question of the measure. Internal reliability was assessed using Cronbach's Alpha. In evaluating validity we examined the directions and levels of correlations between the ASCOT measures and our other measures of resident experience and quality.

The qualitative analysis, which was conducted by a researcher not involved in the earlier stages of the research, focused on the textual evidence provided for the ratings, which fieldworkers were requested to make for each of the individual

SCRQOL domains with and without services. The analysis examined the reliability of the buddy observations and individual fieldworkers, and examined any difference or similarity in the basis for the ratings in LD and OA homes. Appendix E shows examples of the type of evidence provided against ratings of experienced and expected SCRQOL in the absence of services for each domain.

Nvivo was used to organise the qualitative information on fieldworkers' rationales of the ASCOT rating in a more readable, transcription type format. These were coded to group key domains (Control, Meals, and so on). These domains were then split into individual files to isolate the rating and observer notes from the rest of the data. This allowed for a more unbiased analysis of the domain data. For some comparative aspects of the analysis, Excel spreadsheets were created from the SPSS quantitative data sets and used to look at the quantitative data in a tabular format and to include the buddy ratings provided. Field notes for the SCRQOL section of the toolkit for the buddy and individual care homes were also reviewed. The researcher interviewed the other team members involved in the fieldwork management to check ideas and obtain background on the project, the training experience and tools used, information gathering process, observational raters and care home issues.

To analyse variations in the basis of fieldworkers' judgements:

- An assessment was made of their observational style, by reviewing their notes without connection to ratings or background information on the raters themselves.
- Information was gathered from team members who worked with the fieldworkers to get a sense of bias, strengths and weaknesses. This was then compared with the independent assessment.
- The data covering ratings and evidence notes were reviewed to get a sense of ratings that the observers made in relation to the evidence provided by them. Any problems with ratings generally were more obvious here.
- The data covering the buddy homes were reviewed to assess how grouped raters performed the same observation. What became more obvious here were indications of how strong the supporting evidence arguments were in relation to the ratings and the extent to which one observational rater of a pair appeared stronger than the other, usually because that rater provided more detailed evidence to support their assessments than the other rater.

- The individual (non-buddy) home observations were reviewed to see if the impressions of how reliable the raters were in the buddy homes were similar to how they performed on their own.

4.2 Inter-rater reliability and SCRQOL ratings

Table 4.1 shows the percentage agreement between ratings and Kappa values for each domain and average overall for the unweighted measure. Three sets of findings are shown. The first columns on the overall sample include all the data for all the buddy observations made. The second, ‘wave 2’ sample only includes those observations at the second wave, after people had more experience in rating. The third column reports on a reduced sample over both waves. Qualitative analysis (see below) identified that one particular fieldworker’s ratings were particularly poor, and notes were very sketchy. One service was also identified as a problem where the first time the observations were undertaken the fieldworkers completed this section together, having misunderstood the instructions. The reduced sample excluded these services and also lost the distinction between ‘no needs’ and ‘all needs met’ in the rating of current SCRQOL. This distinction has not been maintained in the measure used in the subsequent analyses, as there is no separate preference weighting available.

Table 4.1: Inter-rater reliability of current and expected SCRQOL

	Overall (n=113)		Wave 2 (n=59)		Overall reduced sample (n=95)	
	Kappa	% agreement	Kappa	% agreement	Kappa	% agreement
Control						
Current	0.33	65	0.36	65	0.36	67
Expected	0.44	3	0.35	69	0.40	71
Personal care						
Current	0.42	78	0.37	81	0.50	85
Expected	0.66	83	0.72	87	0.73	86
Meals						
Current	0.34	65	0.25	61	0.46	72
Expected	0.54	81	0.44	78	0.58	82
Safety						
Current	0.41	86	0.36	83	0.52	88
Expected	0.41	74	0.47	80	0.45	76
Social participation						
Current	0.39	67	0.33	65	0.51	74
Expected	0.47	75	0.43	72	0.55	78
Occupation						
Current	0.33	63	0.40	69	0.39	66
Expected	0.60	81	0.59	81	0.65	83
Accommodation						
Current	0.48	83	0.40	74	0.58	88
Expected	0.64	85	0.65	85	0.73	88
Mean						
Current	0.39	72	0.35	71	0.47	77
Expected	0.52	78	0.53	79	0.57	81

Generally, ratings in the absence of services appear to be more reliable and therefore potentially easier to complete. Although the Kappa statistics are below the normally acceptable level of 0.6 for the ratings with service (average Kappa 0.53), the percentage agreement almost reaches the generally acceptable level for high agreement of 0.8. The domain of control appears to be the hardest to rate ,at least in the sense that it gives rise to more disagreements between fieldworkers.

Excluding the initial ratings improved levels of agreement for some domains, but the picture was not consistent. However, the reduced sample does show a higher level of reliability. This was primarily due to the exclusion of the fieldworker that provided inadequate evidence. Ratings from this fieldworker were excluded in subsequent analyses. The reduction of number of levels of response (from four to three) had a minimal effect, leading to a slight increase in reliability for some items but not for others.

Examining the individual items to identify where disagreements lay showed that where there were differences these were usually between ratings of levels 1 and 2 or between levels 2 and 3. Wider disagreements were rare: none in the domains of Social participation, Occupation, Accommodation or Meals, only one in the domains of Control and Safety and two in the domain of Personal care.

The unweighted total measure was estimated and agreement between observers tested on this using Spearman's rank order correlations on the total scores for the reduced sample. There were significant correlations between observers for both the total with service ($r = 0.618$ $p<0.001$) and in the absence of service ($r = 0.723$ $p<0.001$).

The qualitative analysis provided helpful insights into what lay behind these findings. This indicated that there were some buddy pairs that provided a strong level of reliability in that the ratings and evidence was balanced and appropriate and individualised to the two observers. There was a middle range, which included most fieldworkers, where one observational rater appeared stronger than the other. The ratings may be slightly different, but one rater was able to more strongly support their ratings through detailed and measured evidence. The impression was of that rater being more reliable in assessing the situation at hand. Overall, the raters agreed about which situation was better or worse, even if there were slight differences.

However, there were some pairings that provided very poor evidence, the ratings seemed unbalanced or the ratings were the same across the board (i.e. restricted to 2/3 for all domains and individuals observed). This raised the concern that the ratings were either overstated or understated or that time and care were not taken to assess the situations sufficiently. In some cases there were different levels of consistency in different homes visited by the same buddy pairs. This might be because the situation observed at some homes was more open to individual assessment than others. For example, some buddy pairings were rating in a similar style in some homes and then in another home it appeared as if they were observing a different person. This was checked and there was no evidence that a different person was observed.

In terms of the types of evidence provided, there were a few fieldworkers who were performing in an automated way. They were quite mechanical, consistent, and supported their choices in a sound and balanced manner. For the greater group of raters, the reliability was more that of independent witnesses. They demonstrated their independence by disagreeing slightly, but were behaving with good judgement and supporting their evidence in a balanced way with some slight variations. There was a small group that provided very limited and not entirely consistent evidence, who may not have had the skills needed, were thrown off by the reality of the observation process or just require more experience and training.

4.3 Individual rater reliability

Many of those fieldworkers who performed well in the buddy assessments seemed even more descriptive in their notes when making ratings individually in homes, and emerged as quite solid performers individually. By 'performance' we mean that the ratings were fair, based on the notes provided and the notes were descriptive enough to support them. The rating and notes were varied, providing evidence that a thought process was going on and that ratings were not arbitrarily applied. One individual who appeared to be poor in the buddy assessments, had been identified as a good performer during training. When individual assessments were reviewed, their ratings were much better, in terms of range and evidence provided, than those made during the buddy pairings.

There were a small group of observers who did not perform well. As we identify above, the decision was made to pull the ratings of one fieldworker from the main analysis as early on in the qualitative analysis it became evident that she provided extremely basic notes or did not give any evidence at all. This person also had a poor performing buddy, who performed better when making individual assessments. Another rater had a tendency to use the same limited range of ratings all the time. This did continue into the individual assessments and the notes were scant at times. There were two observers whose performances were very variable; they both had occasions where they did very well and the ratings and notes seemed sound, and then other occasions where the supporting notes were very basic or scant. It is possible that these observers were generally performing well but had the occasional difficult day where there was a stressful or difficult environment in which to observe.

For the most part, other than these few specific issues, the evidence suggested that judgements were being made on a reliable basis, with two observers coding in an automated and consistent way, suggesting that future training could draw on this to improve the ratings in future.

4.4 The ‘Dignity’ domain

The domain of Dignity is intended to reflect the way that staff treated the resident, picking up those elements of the care process that would only be expected to affect the quality of life of people in receipt of care and support, rather than those aspects of process that directly affect our domains of control, personal cleanliness and so on, which are equally applicable to people without care needs. Conceptual work undertaken after the fieldwork commenced, as part of a separate project (Netten et al., 2009), has identified this is best defined in terms of the impact on self-esteem: how the person feels as a result of the way they are treated when being provided with care and support.

The quantitative analysis of inter-rater reliability on the Dignity domain was not obviously much poorer than some of the other domains. Collapsing the ratings to three levels, as was done for the other domains, Kappa was 0.38 and the percentage agreement was 97. For 91 people, both observers rated Dignity as no needs, for one person both observers rated as low needs. This meant they disagreed on only three cases in the reduced sample, all between no needs and low needs.

However, the qualitative analysis of the basis of the ratings raised some concerns. Evidence provided by the fieldworkers was quite basic and tended to indicate more ‘subjective’ views. In the other seven domains the notes demonstrated a similar focus among the fieldworkers; this was much less evident for the Dignity domain. Training had provided detailed and clear examples of poor

Dignity, but very positive examples for the Dignity domain were not identified. The resulting ratings suggested a very basic assessment of Dignity – only identifying very poor practice.

As a result of this analysis it was decided not to include the Dignity domain in the subsequent measure.

4.5 Rating OA and LD residents

We reviewed and compared the types and level of evidence provided for OA and people with LD, taking into account the previous training and experience of the fieldworkers. The level of background knowledge of the observer did affect the amount and type of evidence provided. When an observer with a background in LD services was reporting on residents with LD, the notes were more descriptive.

Generally, however, more detailed observational evidence was cited when there was severe cognitive impairment, be it as a result of learning disabilities or dementia (or both). This may reflect the fact that for these groups, where it would be less likely to be possible to interview them to get their perspective on current SCRQOL, more reliance was put on observational data.

In terms of expected needs in the absence of services, supporting evidence was more detailed when the individual had a diagnosed condition which had an expected pathway or milestones in terms of the severity of the condition. Usually this was in the case of old age dementia, Parkinson’s or LD-specific conditions. Ratings of expected needs were also influenced by the attitude of the residents to the service. Several of the fieldworkers identified that OA and LD had different attitudes to activities and support provided by the home. While LD residents were encouraged and generally welcomed undertaking as many everyday tasks as possible, sometimes OA residents were quite able but did not want to participate in what they saw as mundane tasks. This has been found in previous research, where older residents have reported that, if they were living in a home, it was time ‘to put their feet up and have someone else do the tea’ (Towers, 2006). This does not mean that they did not want to be occupied; more that they did not want to participate in these types of task.

4.6 Internal reliability of ASCOT measure

To evaluate internal consistency we used the unweighted scales (which assume that high level needs score 0, low level needs 0.5 and no needs 1) reflecting current and expected SCRQOL in the absence of services. The internal consistency of the measure was good and, as with reliability, notably better for the ‘expected’ measure than that reflecting current levels of SCRQOL. For the whole sample Cronbach’s alpha for current SCRQOL was 0.71 and for expected SCRQOL 0.95. For the buddy homes internal reliability was very similar for the two sets of observers: for current SCRQOL 0.65 and 0.68; for expected SCRQOL 0.93 and 0.92.

There is very little room for improvement for the ‘expected’ scale. For current SCRQOL the domains which had the lowest correlations were Safety and Accommodation. This may partly be to do with distribution of ratings. Most people had safety rated as being all needs met, so when there was a disagreement as such it had a bigger effect on the reliability. Similarly for accommodation: in general, homes were not observed as dirty and badly kept.

4.7 Reliability of other measures

While our principal focus is on the ASCOT measures of SCRQOL, it is important to understand the reliability of the other measures used (such as the Active Support Measure and Nursing Home Quality Scale) as these both fed into the ratings of the current and expected SCRQOL of our focal residents and were explored in the analyses. Table 4.2 summarises the reliability statistics for each of the other sections of the observational toolkit. For the section on pain, reliability was not calculated for each individual item as so few examples were observed: there were only 16 people for whom pain was noted by either observer. Agreement on whether pain was observed or not was at 90 per cent for total reliability (Kappa=0.44).

Table 4.2: Inter-rater reliability of observational toolkit sections

	n	Kappa mean (range) all questions/ categories	Percentage agreement (total)	Spearman's correlation co-efficient (mean)
Section 1 Engagement in meaningful activity and relationships (EMACR)	3470 minutes of observations, 112 people in 28 services	0.60 (0.5 – 0.80)	94 (88-99)	n/a
Section 2 Active support measure (ASM)	112 people in 28 services	0.50 (0.3 – 0.7)	70 (56 – 100)	0.60 (0.30 – 0.71)
Choice and communication	112 people in 28 services	0.51 (0.49 – 0.53)	73 (71-74%)	0.60 (0.59 – 0.62)
Section 3 Environment	28 services	0.631 (0.47 – 0.78)	86 (75-96)	
Homelikeness	28 services	0.4 (0.2 – 0.5)	53 (48-64%)	0.66
Section 4 Nursing home quality scale (NHQS)	28 services	0.35 (0.12 – 0.76)	55 (33-91)	n/a
Section 5 Health and nutrition	28 services	0.76	92	n/a
Exercise	113 people	0.59	93	n/a

Reliability of measures at the individual level was generally good, with average Kappa statistic values around 0.6 and percentage agreement over 0.8. For the Active Support Measure, this was slightly lower, although percentage agreement could still be considered as acceptable (0.7 and above). The Active Support Measure is a rating scale that requires a more subjective judgement of what is happening for people, and requires some experience of the type of support that would be rated highly by the scale. While some of the fieldworkers had that experience, it is possible that the training, while intensive, was not sufficient to embed that knowledge in those who did not already have it. It is also the case that quality of support is difficult to rate if people are not engaged in activities, and staff are not around to support people: this was the case in many services in this sample. Reliability was poorer on the measures which were rated at service level: for example, on the NHQS. However this is likely to reflect the limited number of observations at the home level (28), the relatively large number of items (36) and that many of these are six point items. Similarly, homelikeness seemed difficult to rate reliably. Again, this was a five-point scale with only 28 cases.

4.8 Validity of ASCOT

While it was not possible to formally test the validity of the ASCOT measures, we could examine the association between the SCRQOL domain scores and other measures to see if the directions of effect were what we would expect them to be.

There were significant Spearman’s rank order correlation coefficients ($p<0.001$) between all domains on the ASCOT measures and the Active Support Measure percent score. The strongest relationships were between Meals with service ($r=-0.360$), Occupation with services ($r=-0.351$), and Control with services ($r=-0.291$). Percentage of time engaged in activities was also correlated with most domains of the ASCOT measure (the only exceptions being Safety with service and Accommodation with service). The strongest correlations were with Control in the absence of services ($r=-0.314$), Personal care in the absence of services ($r=-0.360$), and Accommodation in the absence of services ($r=-0.308$).

There were also correlations between the ASCOT domain scores and the scores on the Index of Community Involvement (ICI), the Index in Participation in Daily Life (IPDL) and on the choice-making scale (CMS): as for engagement these were strongest for the domains in the absence of services. This is likely to be a proxy for level of ability: so those who are more able are generally more engaged and participating more and therefore more likely to be judged as able to do more for themselves, more in control and so on.

There were correlations ($p<0.001$) between the total score on the Nursing Home quality scale and Control with service ($r=0.354$), Personal care with service ($r=0.172$), Meals with service ($r=0.236$), Safety with service ($r=0.215$), Social participation with service ($r=0.210$) and Dignity ($r=0.423$). There were also significant correlations between individual items on the NHQS and the with service ASCOT domains. Although, as we identify above, we have reservations about the Dignity domain, it was interesting to note that ratings of dignity were most strongly correlated with NHQS domains (Conversations friendly, $r=0.400$; Staff use resident's name, $r=0.261$; staff acknowledge residents, $r=0.330$; positive interactions, $r=0.286$; Staff appear caring, $r=0.395$; staff treat people with respect $r=0.463$; dressed and clean $r=0.211$). Given the nature of the questions in the NHQS, it is not surprising that it is more closely linked with current needs and dignity than with needs in the absence of service.

4.9 Conclusions

The focus of the study is on finding ways to measure and monitor social care outcomes for a challenging group of people: residents of OA and LD homes. In order to have a consistent basis for measurement, we constructed a complex toolkit, and provided guidance and training in both observation and making judgements on current and expected SCRQOL in the absence of services. Inter-rater reliability is not as high as we might have liked, although, given the challenges presented by the task, perhaps this was to be expected on the first full application of the approach. The qualitative analysis provided helpful insights into what lay behind this, allowing us to make adjustments and allowances in our subsequent analyses (see Chapters 5 and 7). The results also help us identify how the approach might be improved in the future through additional guidance (drawing on best practice among the fieldworkers) and improved training. Concerns were raised about the evidence base for the Dignity domain, which is an area for future development. We return to the potential way forward in Chapter 8.

5. Dependency and capacity to benefit

A fundamental aspect of the proposed approach to measuring and reflecting quality-weighted outputs is the concept of capacity for benefit (CfB). As we explain in Chapter 1, this reflects what a service could potentially provide and reflects both the number of people served and their capacity *to* benefit (CtB) from the service. This latter measure is based on people's expected needs in the absence of the care (expected SCRQOL) and ideal maximum level of SCRQOL in each domain. Our observational approach designed to capture this has been described in previous chapters. Here we link this resource-intensive approach to more easily established measures that might be used routinely to reflect changing CfB of care homes over time.

We start by identifying, for our sample of focal residents, the levels of expected need (expected SCRQOL) in our domains in the absence of the care and support provided, and the assumptions underlying CtB and CfB. We describe models that predict CtB for older adults and people with learning disabilities and discuss what may lie behind the results.

5.1 Expected needs and capacity to benefit

Table 5.1 and Figures 5.1 and 5.2 show the proportions of older people and people with learning disabilities (LD) who had high, low and no expected needs in the absence of the service.

On the whole, older adults (OA) in the sample tended to have more high-level needs, particularly in basic self-care aspects of SCRQOL, than the residents of homes for people with LD, although in terms of safety the picture was very similar, and slightly more needs were identified for the people with LD in the domains of Social Participation and Occupation.

Table 5.1: Expected needs in the absence of care and support

	Older adults **%**	**Learning disabled** **%**	**Whole sample** **%**
Control	*n=372*	*n=364*	*n=736*
High needs	51.1	46.4	48.8
Low needs	43.3	50.8	47
No needs	5.6	2.7	4.2
Personal care	*n=370*	*n=364*	*N=734*
High needs	57.0	45.3	51.2
Low needs	37.0	47.5	42.2
No needs	5.9	7.1	6.5
Meals	*n=372*	*n=363*	*N=735*
High needs	64.0	53.7	58.9
Low needs	33.6	42.1	37.8
No needs	2.4	4.1	3.3
Safety	*n=372*	*n=364*	*N=736*
High needs	60.5	60.2	60.3
Low needs	37.4	38.2	37.8
No needs	2.2	1.6	1.9
Social Participation	*n=370*	*n=364*	*N=734*
High needs	50.0	56.9	53.4
Low needs	44.6	40.9	42.8
No needs	5.4	2.2	3.8
Occupation	*n=372*	*n=363*	*N=735*
High needs	56.2	57.9	57
Low needs	39.0	40.2	39.6
No needs	4.8	1.9	3.4
Accommodation	*n=369*	*n=361*	*N=730*
High needs	67.5	60.7	64.1
Low needs	30.1	36.0	33
No needs	2.4	3.3	2.9

Figure 5.1: Older adults' expected social care-related quality of life in the absence of services

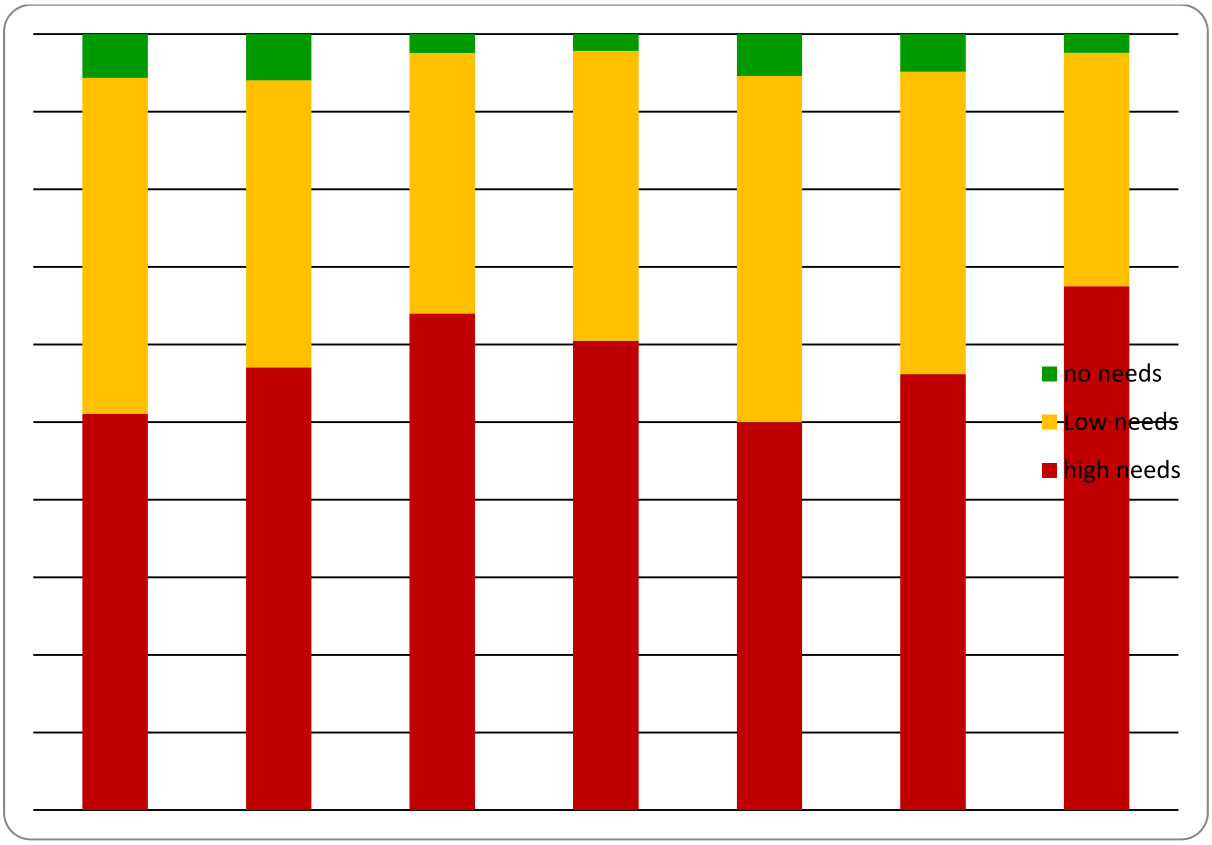
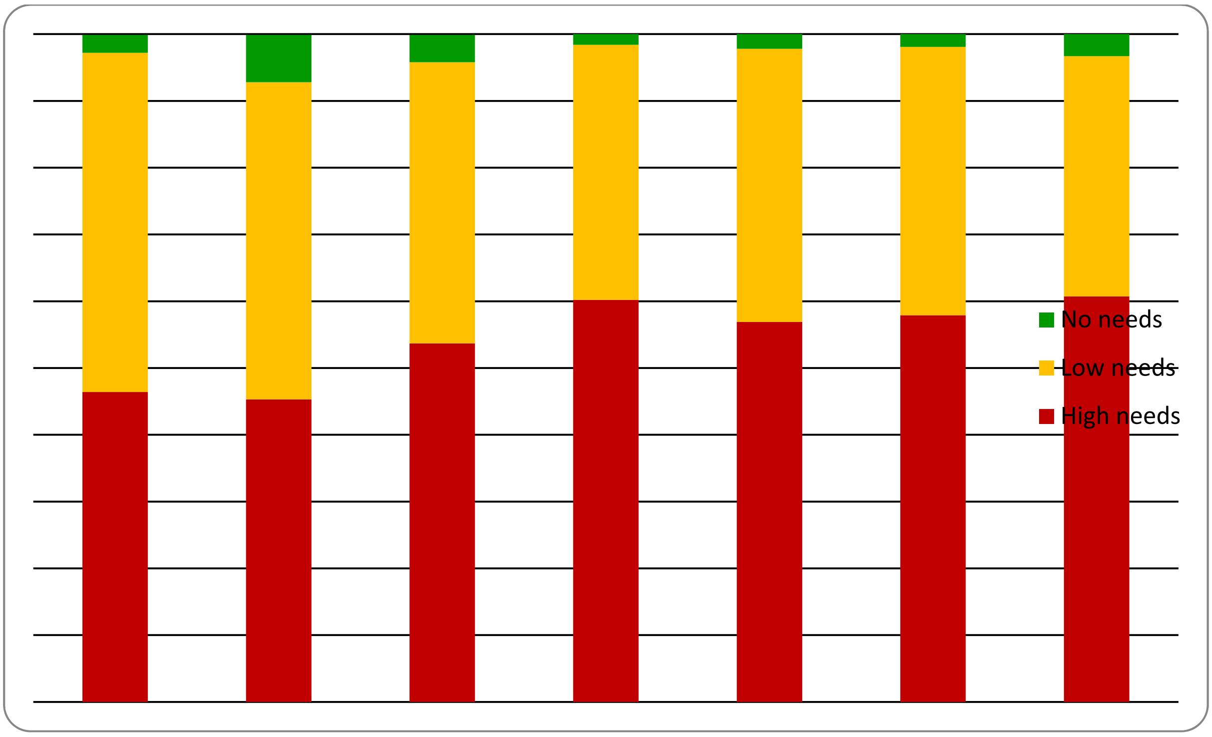


Figure 5.2: People with LD expected social care-related quality of life in the absence of services



In creating a composite measure we want to reflect the relative importance of the different aspects of SCRQOL. To do this we incorporate population preference weights (see Appendix A) and standardise the resulting score so it lies between 0 and 1. The average overall level of expected social care related quality of life (SCRQOL) in the absence of services was not significantly different for OA (mean 0.26, median 0.18, range 0-1) than for people with LD (mean 0.28, median 0.29, range 0 - 1). However, the distribution of expected SQRQOL is highly skewed with over a third of OA and just under third of people with LD scoring 0, indicating high level needs in all domains in the absence of the service (see Figures 5.3 and 5.4).

Figure 5.3: OA expected SQRQOL in the absence of services

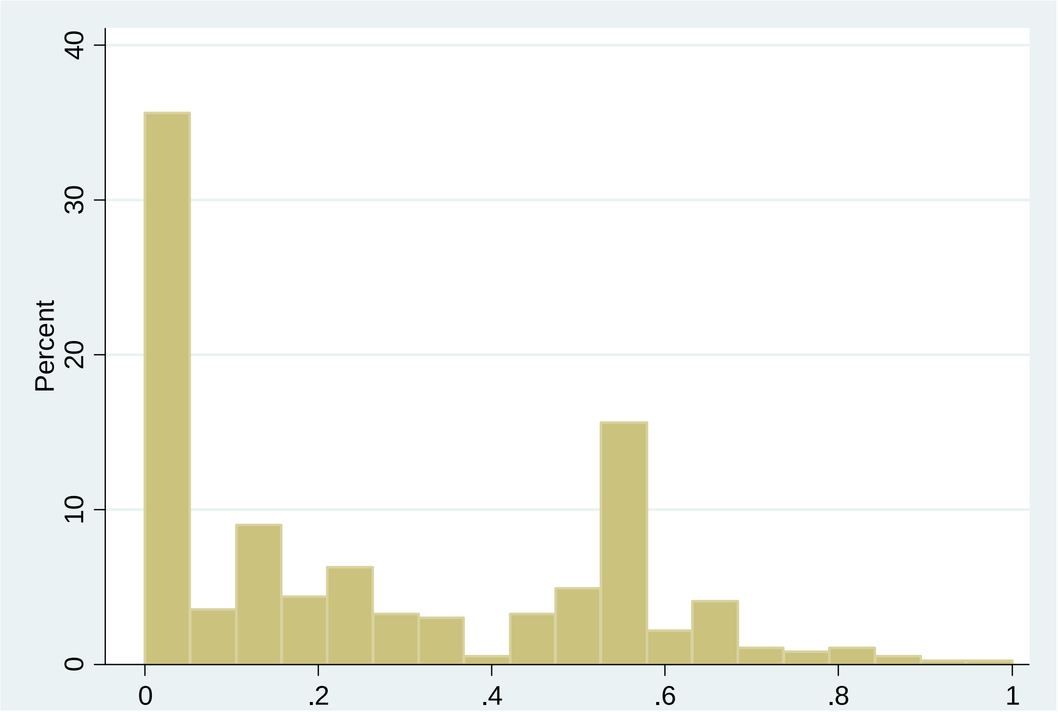
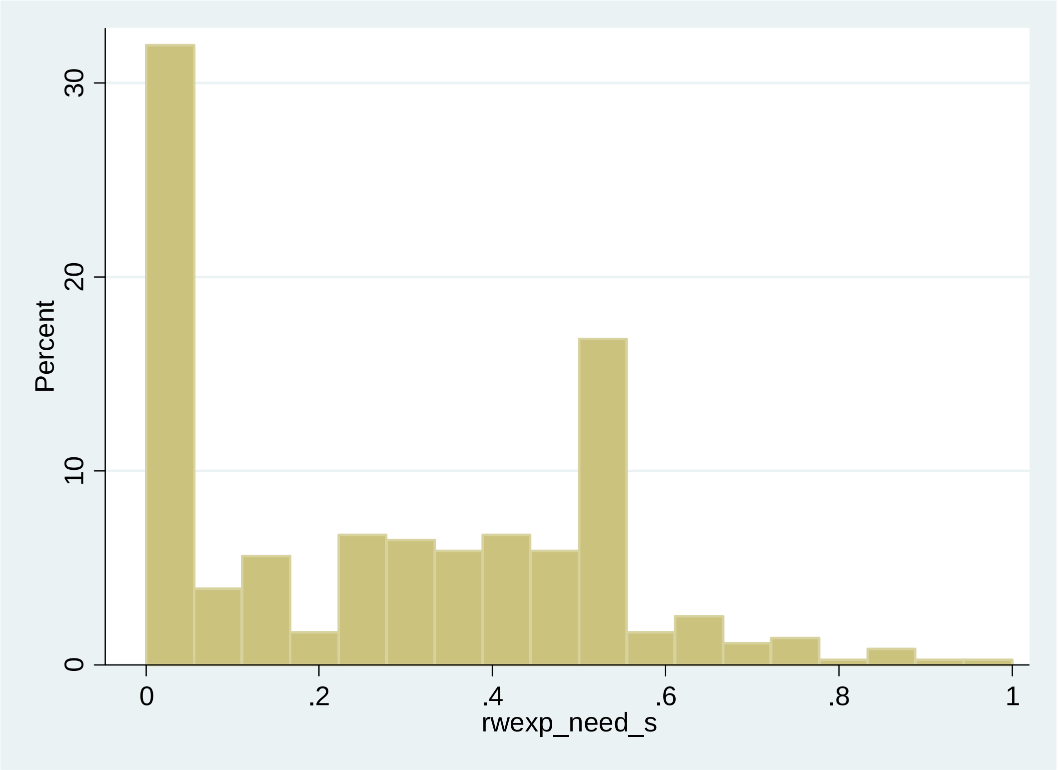


Figure 5.4: LD expected SQRQOL in the absence of services



As care homes are concerned with all aspects of social care related quality of life, the difference between this expected score and the maximum SCRQOL score : where all needs are met in all domains - is that individual's CtB from the service. There were no significant differences between older adults and people with learning disability in terms of their capacity to benefit : the average CtB scores for our sample were 0.74 (SD 0.25, 0 – 1) for OA and 0.72 (SD 0.25, 0 – 1) for people with LD.

Table 5.2 shows the mean scores of residents' CtB by the type of home and sector. As we would expect, CtB is significantly higher in nursing homes ($p<0.001$) overall and for OA homes, although the table shows that CtB was relatively low for the four residents of the one LD nursing home. There was a significant difference between sector among residents of OA homes ($p = 0.03$), with residents of voluntary homes showing lowest CtB, but no significant difference for LD homes ($p = 0.60$) or the sample as a whole.

Table 5.2: Mean Capacity to Benefit scores by home type

	Older people		Younger Adults		All	
	Mean	N	Mean	N	Mean	N
Type of Home						
Personal care only	0.69	217	0.73	330	0.72	547
Nursing	0.81	134	0.56	4	0.80	138
Sector						
Private	0.75	275	0.71	191	0.74	466
Voluntary	0.67	62	0.72	127	0.70	189
LA/NHS/Other	0.81	23	0.76	39	0.78	62
All	0.74	365	0.72	357	0.731	722

5.2 Predicting CfB and CtB

In order to monitor how the capacity for benefit (CfB) of a service changes over time, we need to be able to predict the average capacity to benefit (CtB) of the individuals using either indicators that are already in use routinely or low-burden measures that could be introduced. The first step in this process is to identify the best model for predicting CtB. We would expect CtB to be associated with other measures of dependency such as ADLs and possibly other individual-level factors, such as age and gender. In exploring the relationship with other indicators of dependency, we drew on the full range of dependency measures described in Chapter 2, including the EQ5D, which is more usually described as a health outcome measure, to identify the best indicators to predict individual CtB.[12]

Our sample of residents is clustered in the homes so may share characteristics – perhaps as a result of being placed in homes with residents with similar characteristics. As a result CtB might also be associated with home level characteristics, so to avoid misspecifying the model we need to allow for this data structure in our estimates. We tested to see if this was the case and initial diagnostic analyses identified that the intraclass correlation coefficients for both samples (OA, LD) were highly significant, suggesting that multi-level modelling was required. The form of the models estimated was:

---

[12] Appendix F shows the distributions of the variables included in the final model.

$$Y_{ij} = \beta_{00} + \sum_{g=1}^{s} \beta_g x_{gij} + \sum_{g=1}^{t} \gamma_g h_{gj} + e_{ij} + u_{0j}$$

Where:

Y_{ij} is CtB$_{ij}$ - capacity to benefit of the *i*'th resident in the *j*'th home

β_{00} is the average intercept

x_{ij} are the characteristics of the *i*'th resident in the *j*'th home.

h_j are home level characteristics

e_{ij} reflects the random effect at resident level

u_{0j} reflects the random effect at home level

Both Tobit and Generalised least squares (GLS) models were run with random effects at the individual and home level. Tobit allows for the 'censored' nature of the data: we only reflected whether there were any needs in our SCRQOL domains as the measure stood at the time of the study.[13] In practice, in terms of 'expected needs' censoring was most relevant at high level needs with, as we describe above, a substantial proportion of residents identified as having the very highest needs in each domain. The estimations, nonetheless, produced very similar results and therefore, on grounds of parsimony, we report the GLS models.

Tables 5.3 and 5.4 show the models that best explained OA and LD home residents' CtB respectively. The models had good explanatory power, with R^2 in the region of 0.70 in both cases. As we note on the table, in order to allow for variations in ratings between fieldworkers we included dummies for each of the fieldworkers, selecting the reference fieldworker at random. Fieldworker effects accounted for 11 per cent of the total variance in both models, thereby contributing to explained variance at the home level. Chapter 4 discusses what lies behind this variation.

---

[13] Further work is developing a four level version of our measure that reflects variations in 'capabilities' or more aspirational states beyond having 'no needs' or 'needs met' in each domain (Netten et al., 2009).

Table 5.3: Model[14] of OA residents' capacity to benefit

	Coefficient	Probability
Resident characteristics		
Needs help dressing	0.77	<0.001
Needs help with personal affairs	0.41	0.001
Ability to care for him-/herself (EQ5D self care)		
(cf. no problems with self care)		
Some probs with wash/dressing - all providers	0.75	<0.001*
Some probs with wash/dressing - Vol provider	-0.03	0.906*
Unable to wash/dress - all providers	0.93	<0.001*
Unable to wash/dress - vol providers	0.47	0.159*
Not anxious or depressed (EQ5D-item)	-0.23	0.01
Communication (MDS CPS-item) (cf. understood)		
Usually understood	0.16	0.191
Sometimes understood	0.48	<0.001
Rarely/never understood	0.56	<0.001
Home characteristics		
Nursing home Care Quality (NHCQ)		
NHQS process	-0.01	0.011
NHQS structure	0.03	0.16
Home type (cf. personal care only)		
Nursing home	-0.15	0.226
Home sector (cf. all non-voluntary sector)		
Voluntary sector	-0.20	0.095*
Constant	2.54	
Random effect		
Level 2 (care homes) variance	0.15	
Level 1 (residents) variance	0.66	

* Derived from bootstrapped standard errors[15]; R^2 = 0.70 (overall), 0.80 (between), 0.56 (within); Wald chi square statistic =525.7, p<0.001

---

[14] Although co-efficients not shown model included fieldworker dummies.

[15] Bootstrapping is a method for estimating the standard errors of variables when they interact with other variables e.g. the impact of star rating by sector. It involves repeated re-sampling of the data and re-estimation of the regression model and calculation of the marginal effects (of the interactions) on each re-

Table 5.4: Model[16] of LD residents' capacity to benefit

	Coefficient	Probability
Resident characteristics		
Number of ADL problems (0 to 3)[1]	0.16	0.004
Self-care ability (EQ5D-item)		
(cf. no problems with self care)		
Some problems with wash/dressing	0.34	0.006
Unable to wash/dress	0.64	<0.001
Mobility (EQ5D-item) (cf. no problems)		
Problems walking or confined to bed	0.33	<0.001
Communication (MDS CPS) (cf. understood)		
Usually understood	0.13	0.265
Sometimes understood	0.65	<0.001
Rarely/never understood	1.06	<0.001
Home characteristics[5]		
Specialist home (cf. non-specialist)	-0.42	0.005
Star rating (cf. zero or one-star)		
Two-star	0.17	0.158
Three-star	0.48	0.004
Constant	1.95	
Random effect		
Level 2 (care homes) variance	0.29	
Level 1 (residents) variance	0.63	

[1] Three ADLs are: walking indoors, getting in and out of bed and/or wash face and hands; R^2 = 0.65 (overall), 0.79 (between), 0.43 (within); Wald chi square statistic =464.3, p<0.001

---

sample. In this case we re-sampled 1000 times. The distribution of coefficient and marginal effect estimations for these 1000 re-samples is then used to estimate the associated standard error and hence judge statistical significance.

[16] Although coefficients not shown, model included fieldworker dummies.

While there is some variation in the specific measures that work best in predicting CtB at an individual resident level, there is a lot of overlap between the model for older adults and people with learning disabilities. Abilities to wash, dress and communicate are clearly very important. For people with LD, mobility is also an issue, probably reflecting the fact that mobility problems are less widespread in this population than OA, so has more of an additional impact on their CtB. Similarly, for OA needing help with personal affairs, something likely to be associated with both cognitive impairment and severe frailty, is associated with CtB, whereas there was much less variation in this among people with LD in our sample (see Appendix D). Anxiety and depression also have an impact on the CtB of OA but were less of an issue among people with LD.

In addition to the variables shown in the final LD model, there were some indications that family contact had implications for residents' CtB from services. In one model, fortnightly visits compared with less frequent visits appeared to reduce CtB from the service: the direction of effect we would expect if families were contributing to meeting residents' needs. However, higher levels of contact were not associated with any reduction, so the variable was left out of the final model as an inconsistent indicator of any such effects.

Although our primary interest is in identifying individual level variables that can be used to predict CtB, as we identify above, the data structure requires the inclusion of home level characteristics for a correctly specified model. These characteristics can be held constant (at their mean level) when using the model to predict individual residents' CtB, but if we omit them we are in danger of biasing our estimates. While the nature of the sampling process meant it was not surprising that there was a home level effect, it was a strong effect and it is of interest to consider what lies behind it.

There was no statistically significant relationship with registration as a nursing home, which suggests that the higher levels of dependency in these homes were picked up through the individual OA resident-level characteristics. Home type was statistically significant when the model was estimated with individual impairment variables removed. We investigated whether there was a relationship between CtB and home specialism. In homes for people with LD there was a relationship, which remained when individual resident characteristics were included in the model. In about half of these homes the specialisation reflected particular religious or ethnic groups; the other half specialised in terms of type of impairment or condition (e.g sensory impairments, dementia, epilepsy).

The results suggested that voluntary sector homes for OA have somewhat different populations than private or LA run homes. We identified above that CtB of residents varied by sector in OA homes. In the sample used in the model of OA homes (n = 269)[17] the difference between voluntary homes compared with those in other sector homes was 0.37 lower than residents of other types of homes, significant at the 10 per cent level (p = 0.06). Some of this difference is due to different levels of impairment of residents in the different types of home. Controlling for other factors such as disability using the regression analysis reported in Table 5.3, this difference falls to 0.2 but is still significant at the 10 per cent level (p = 0.095).

The modelling showed a statistically significant effect of home sector on the relationship between CtB and impairment measures. The relationship between impairment as measured by self-care ability and the CtB score was strongly significant overall, for all types of home, as we would expect. However, for voluntary sector homes specifically, the relationship between self-care impairment and CtB was less strong, suggesting that voluntary sector homes tend to be associated with people with a somewhat different set of needs. We did not find any significant relationship between home sector and CtB for LD homes, although specialism was highly significant as noted above and, as we identified in Chapter 3, more particularly in homes for people with LD, specialist homes are more likely to be in the third sector.

In theory, the star rating of homes should not have a direct effect on people's CtB. Home-level characteristics are only related to mean CtB scores of residents because different types of home cater for different populations of residents. For older people's homes, star rating was not a factor. However, for LD homes, average CtB differed between homes with different star ratings.

The precise nature of the relationship between CtB and quality was different in OA and LD homes and changed when fieldworker dummies were included. For example, for OA the NHQS process quality indicator was statistically significant

---

[17] Missing values on other variables included in the model mean that this sample is less than our total sample of OA home residents.

when the fieldworker dummies were included but not when the variable was excluded. In homes for LD, on the other hand, the statistical significance of home quality factors was reduced when fieldworker dummies were included: the significant differences between two and three-star homes and of the NHQS structure quality indicator, which were evident when the dummies were not included, disappear.

We discuss the differences in the way that fieldworkers were rating SCRQOL in Chapter 4. The inclusion of dummies for each fieldworker allows for this variation, so we should pick up the underlying relationships between CtB and individual and home level characteristics. If this is so, then in homes for OA, an underlying relationship between CtB and care home process quality becomes evident once fieldworker variations are allowed for. In LD homes, the results suggest that there was an association between the quality of the homes (particularly structural aspects of quality) and fieldworkers rating of them. The inclusion of the dummies reduced this effect.

In addition to the home level, we tested whether there was any ‘area effect’, reflecting placement policies. For LD homes there was a borderline statistically significant effect, suggesting that residents of homes in the West Midlands were more likely to have higher expected needs, thus higher CtB, than other areas.

5.3 Discussion and conclusions

We measured expected needs in the absence of the service in order to identify residents’ capacity to benefit from the services. Once we have a means of predicting CtB for individuals, we can predict CfB for homes and link this to data that could be routinely collected over time. In this process, home level factors can be held constant. Our analyses have allowed for variations between fieldworkers in making their judgements, which we identified in Chapter 4 as something that needs to be borne in mind in interpreting the data.

The evidence of expected needs in the absence of the services provided by the homes identifies that they are providing care and support for a very dependent population, with substantial proportions having high-level needs in all domains. These needs are, as we would expect, related to more traditional measures of dependency, with ADL indicators, the EQ5D and the MDS CPS providing useful predictors for both OA and people with LD. The influence of home-level factors was rather less expected. We would anticipate that residents of nursing homes for OA would demonstrate higher CtB, as indeed they do. However, the impact of quality and sector was unexpected and interesting.

The results for LD suggested that even after accounting for resident characteristics, the home-level star rating was picking up some further influence on resident CtB, such that homes rated as excellent have resident populations that have somewhat higher observed CtB (i.e. needs in the absence of services) than homes with other star ratings, other things being equal. For OA the NHQS appeared to pick up a similar effect. This additional CtB might reflect the impact of characteristics such as basic motivation, attitude or personality, resulting in people being ‘difficult’. Such characteristics would not necessarily be classified as challenging behaviour: no behavioural indicators were significantly associated with CtB. However, such residents may be asked to leave some care homes for being disruptive or awkward. We might hypothesise that, by their very nature, the types of homes that can accommodate such people are better quality. Such individuals may have no more severe physical or cognitive impairments, but have higher needs in terms of social participation, control and occupation, for example. The nature of these unmeasured characteristics might be different for OA and people with LD, and fieldworkers may vary in the degree to which they are sensitive to these. It is also possible that in making judgments about needs in the absence of services, fieldworkers are unduly influenced by contextual factors, such as home quality.

We might expect that voluntary sector homes, having a common uniting purpose, might be able to draw on additional input and support from relatives, surrounding communities, volunteers or others. This might explain the impact of the voluntary sector in OA homes and specialism in LD homes, in reducing residents’ capacity to benefit from the services. It may also be that, if these homes fit with people’s beliefs or are seen as particularly trustworthy or expert, people are prepared to move into the homes at an earlier stage than they would be otherwise – again resulting in third sector homes catering for a less dependent population, and thus with lower capacity for benefit. In neither instance should this be seen as the homes providing less valuable services. We need to turn to the outcome generated by the homes to identify the value of what is produced.

6. Outcomes and quality of care

Before we investigate the relationship between the quality ratings of homes and the outcomes for their residents, it is important to have an understanding of people's experiences in the homes, in terms of their current social care-related quality of life (SCRQOL), their outcomes and their day-to-day experienced quality of care and support.

We start by describing the level of residents' current SCRQOL and outcomes, defined as the difference between this and their expected level in the absence of older adults (OA) or learning disability (LD) services. We then describe their day-to-day experiences as observed by the fieldworkers and our measures of quality of care, and describe the association between these and the ratings made by the regulator.

6.1 Current SCRQOL and outcomes

Table 6.1 and Figures 6.1 and 6.2 show the proportions of OA and people with LD who currently had high needs, low needs and no unmet needs, as rated by the fieldworkers.[18] Appendix F shows the equivalent ratings from staff. Just under half (45 per cent) of OA residents and a third (31 per cent) of LD residents were interviewed. Even among those who consented to be interviewed, responses to our structured form of questions were often incomplete. These interviews served more to provide interviewers with further evidence for their observationally-based ratings.

---

[18] Figures 6.1 and 6.2 exclude Dignity as the figures distinguish 'no needs' from 'all needs met', a distinction that is not consistent with our definition of Dignity as the impact of the care and support process.

Table 6.1: Current SCRQOL in each domain

	OA %	LD %	Whole sample %
Control	*n=346*	*n=319*	*n=701*
High needs	4	6	5
Low needs	47	54	51
No needs	49	40	44
Personal care	*n=346*	*n=319*	*n=701*
High needs	1	1	7
Low needs	13	12	13
No needs	86	87	87
Meals	*n=346*	*n=319*	*n=701*
High needs	1	1	7
Low needs	34	31	33
No needs	65	68	67
Safety	*n=345*	*n=319*	*n=700*
High needs	0	0	1
Low needs	14	17	16
No needs	86	83	84
Social participation	*n=346*	*n=319*	*n=701*
High needs	10	6	9
Low needs	47	50	49
No needs	42	44	42
Occupation	*n=346*	*n=319*	*n=701*
High needs	25	11	19
Low needs	46	50	47
No needs	29	38	34
Accommodation	*n=346*	*n=319*	*N=701*
High needs	1	1	1
Low needs	12	15	14
No needs	87	84	85
Dignity	N=346	N=317	N=699
High needs	1	2	1
Low needs	19	34	26
All needs met	80	64	73

Figure 6.1: Older adults' current experienced social care-related quality of life

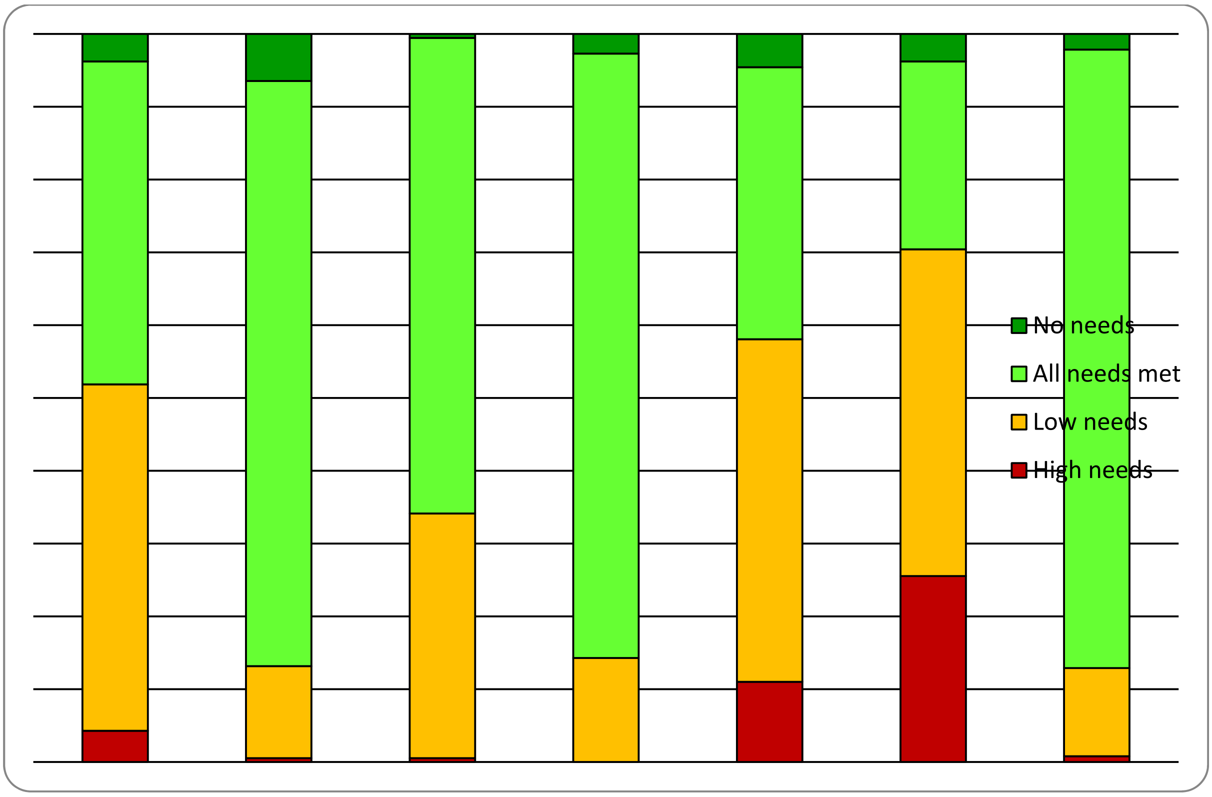
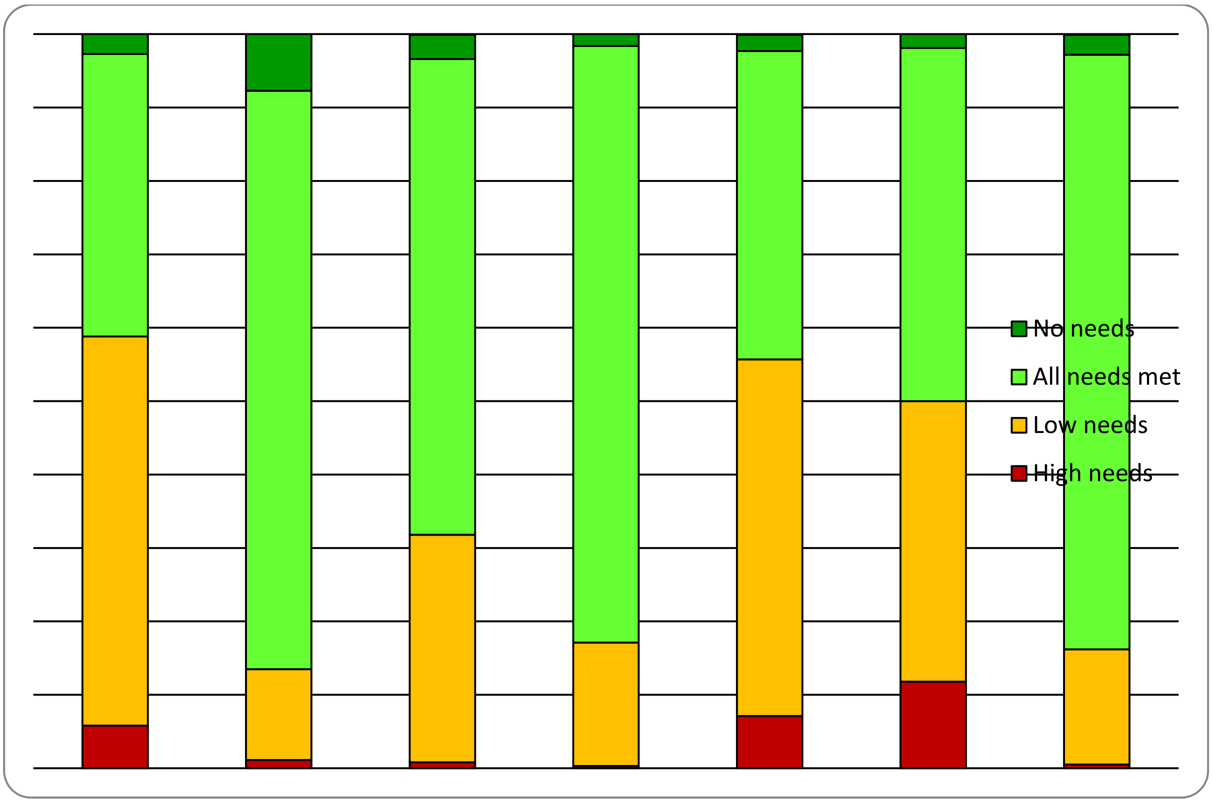


Figure 6.2: People with learning disabilities' current experienced social care-related quality of life



For current SCRQOL we distinguished whether 'no needs' were because the person was managing this aspect of their lives without support from the service (no needs) or whether the needs were being met by the service (all needs met). Reflecting the findings reported in the previous chapter, very few people had 'no needs' to be met by the service. This situation was most likely to apply in the domain of Personal cleanliness and comfort, particularly for people with LD.

There is a very similar pattern overall when OA and LD services are compared, with relatively low levels of unmet needs in fundamental or basic aspects of SCRQOL, such as personal cleanliness, safety, accommodation and meals. However, in the more aspirational areas of people's lives - those we might regard as higher order in terms of Maslov's hierarchy of needs: social participation, control over daily life and occupation - there is considerably more unmet need. This is evident in both types of service but appears to be more marked in homes for OA than LD, with high and low level needs in the domain of occupation for OA particularly prevalent ($p<0.001$). It is also noticeable that, despite living in a communal environment, there is extensive evidence of need among residents in terms of social participation. The Dignity ratings suggest that people in LD homes are more likely to have their dignity undermined by the care process than those in OA homes (p<.001).

Figures 6.3 and 6.4 show the average proportion of the total possible score (using the unweighted summed measure) for each domain in terms of both current and expected level SCRQOL for OA and residents with LD respectively, with higher scores indicating higher quality of life or less needs.[19] The darker shaded area reflects the expected level in the absence of services, and the pale blue area illustrates difference between this and the current experienced SCRQOL. While this is indicative of the level of outcome achieved by the home it does not take into account the relative importance of the domains and levels within them. More basic aspects of SCRQOL are shown on the right of the figures, with the asymmetric pattern of outcome reflecting the difference between these more basic and higher-order aspects of SCRQOL described above.

---

[19] Dignity is excluded as there is no score in the absence of the service.

Figure 6.3: Older adults' outcomes: current and expected SCRQOL

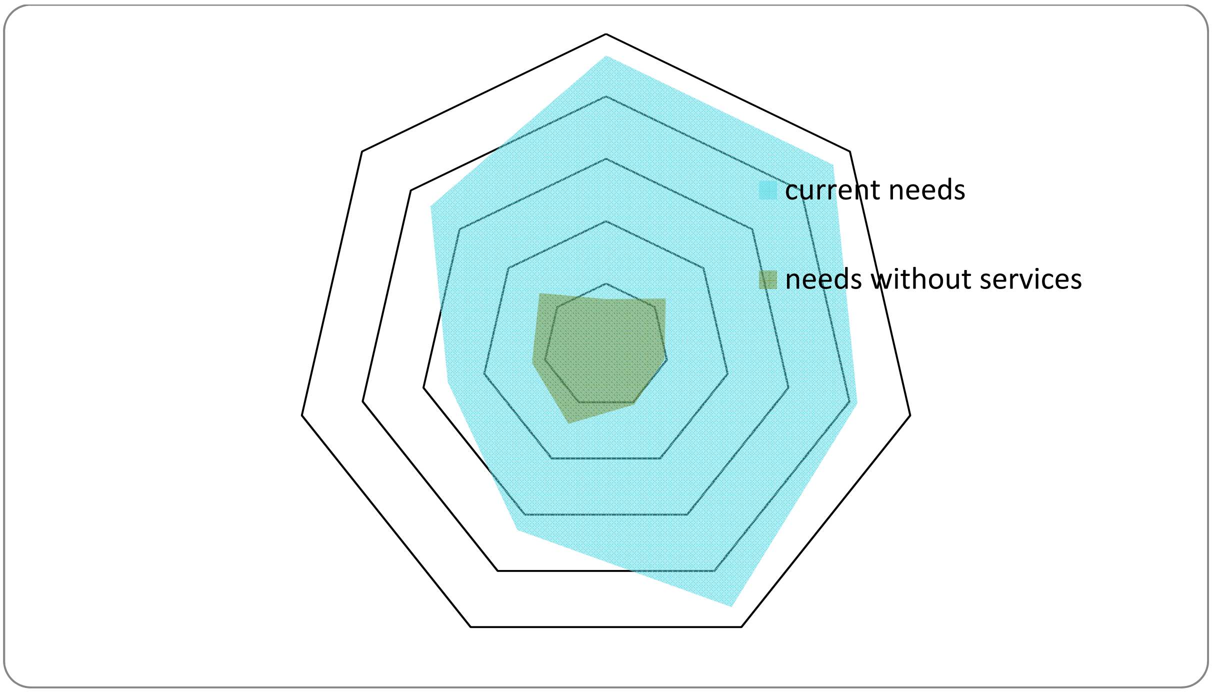
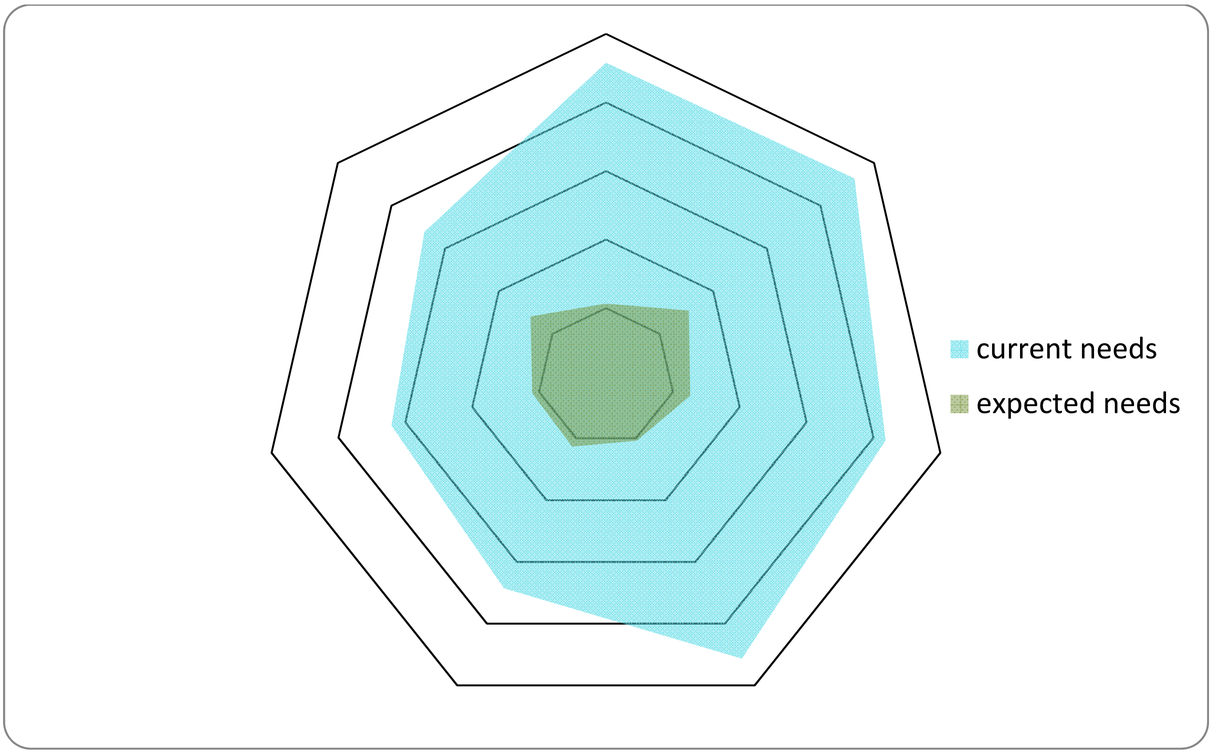


Figure 6.4: People with LDs outcomes: current and expected SCRQOL



While these figures suggest a considerable ‘area’ of outcome, the average scores depicted do not reflect the range of levels of outcome between residents, which were considerable. Some, 4 per cent of OA and 2 per cent of LD residents, were recorded as achieving the maximum possible score: situations where they would have had high level in all domains in the absence of services and scored the maximum possible current SCRQOL. One older adult experienced negative outcomes. Negative outcomes are consistent with situations where people have limited needs and services actually undermine people’s sense of control, social participation or occupation, for example. An example of negative outcome for control is shown in Appendix E.

6.2 Residents’ experiences of engagement, participation and choice

We can gain more insight into the social and occupational aspects of residents’ lives by turning to the results of the detailed observations of residents’ experiences.

The detailed observations took the form of noting what each focal resident was doing and whether they were receiving contact from staff or others at one-minute intervals, rotating around each resident after five minutes to ensure a spread of data across the whole two-hour period. Activities were classified in terms of social or non-social activity, contact and assistance from staff, contact from other residents, and challenging behaviour. At the end of the two-hour observation, the researchers rated the quality of the support received to enable people to take part in meaningful activities and relationships, and also whether staff supported people to make choices and to communicate. Table 6.2 shows the activities in which residents were engaged and Table 6.3 the proportions of time receiving help or other contact from staff, and the quality of the support provided.

Table 6.2: Observed proportion of time spent by activity

	Total sample % (range) (n=680)	**OA % (range)** (n=344)	**LD % (range)** (n=336)
Time spent engaged (overall)	49 (0–100)	48 (0-100)	50 (0-100)
Time engaged in Social activity	11 (0-100)	10 (0–100)	12 (0-78)
Time engaged in non-social activity	39 (0-100)	39 (0-100)	39 (0-100)
Time engaged in repetitive stereotypic behaviour	9 (0-100)	4 (0-92)	15 (0-100)
Engaged <5% of time	8	7	8
Offered no choice during observation	32	25	39
Adapted communication needed but not used	48	43	54
Score on community involvement	33 (0-100)	17 (0-75)	49 (0-100)
Score on participation in tasks of daily living	15 (0–100)	4 (0-100)	27 (0-100)
Score on opportunities for choice making	55 (0-100)	54 (0-100)	56 (0-100)
Observed to take exercise	17	11	23
Observed to do housework	16	5	26
Observed to do gardening	3	2	5
Observed to access out-of-home activity	42	9	72

Table 6.3: Staff support and interaction

	Total sample (n=680)	**OA only** (n=344)	**LD only** (n=336)
% time receiving assistance	3 (0-100)	3 (0-100)	2 (0-25)
% time other staff contact	10 (0-92)	8 (0-92)	12 (0-78)
% Score on active support measure (ASM)	3 (0-95)	35.2 (0-84)	42.8 (0-96)
% no contact from staff	35	45	25
% no assistance from staff	75	80	71
% receiving poor active support (score = <34%)	44	50	37
% receiving good active support (score =>66%)	9	7	13

The data shown in the table reflect the proportion of time spent in each activity out of all the observations made. Observations could not always be made during the period: missed observations were recorded 18 per cent of the time on average (16 per cent for people with LD and 20 per cent for OA). The average number of minutes observed was 24 (ranging from 1-91). Twelve per cent of people were observed for less than 15 minutes (9 per cent of older adults and 15 per cent of people with learning disability). Rerunning the descriptives for those who were observed for at least 15 minutes resulted in slightly lower means but only by 1-3 per cent. Whilst most of the missed observations were accounted for by people being in their rooms with very little contact from anyone, when the door was closed it was not possible to be sure that the person was not engaged in some form of activity: some people may have had the TV or music on and may have been engaged in other things such as self-care activity. On a small number of occasions, ‘missed’ was scored because the person was in the bathroom. It is important to bear in mind that, by taking into account the missing data, the amount of activity may be overestimated, especially in the case of people with learning disabilities who often need staff’s help to be engaged in a meaningful activity.[20]

---

[20] If we do not take account of missing variables and include them in the data, the percentage of time spent engaged reduces from 49 to 39 per cent, and the non-social activity from 39 to 32 per cent.

Although the level of overall engagement for people with learning disabilities was similar to other studies of people of a similar level of functioning, the types of activities observed were almost entirely personal care (eating and drinking, taking medication), and watching TV or listening to music. This was also true of older adults. A few were observed to do drawing or colouring (even though some of them appeared to have a visual impairment) or reading, and a few were observed to take part in simple fetching and carrying-type tasks. A small number of people (mostly those with LD) were observed to prepare snacks or meals and to do household tasks such as loading the dishwasher. The social activity observed was mostly talking to staff, occasionally to other service users. However, the amount of time people received any forms of assistance to engage in meaningful activity was very low: people were receiving help to engage in a social or a non-social activity less than two minutes in every hour. In addition, people were receiving other forms of contact from staff for less than 6 minutes in every hour.

In terms of differences between older adults and people with learning disabilities, there were significance differences in the proportion having no contact from staff (45 per cent for older adults compared with 25 per cent for adults with learning disabilities, $p<0.001$), and repetitive stereotypic behaviours (4 per cent for older adults compared with 15 per cent for adults with learning disabilities, $p<0.001$). In general, people with LD received more contact from staff and were less likely than OA to receive poor-quality support ($p<0.001$). Older adults spent more time engaged in ‘no activity’, but this was accounted for by the fact that those with LD engaged more in stereotypic repetitive behaviour, which was included in the ‘none’ category : this means that older people tended to just sit and stare around them or doze, while people with learning disabilities were more likely to engage in self-stimulatory behaviour when not engaged in a meaningful activity. There was no difference in the percentage of time spent engaged overall during the two-hour observation (48 per cent for older adults and 50 per cent for adults with learning disabilities) but older adults were rated by staff as participating in significantly fewer household and other everyday activities (measured by the Index of Participation in Daily Living) or community-based activities (measured by the Index of Community Involvement) ($p<0.001$). In addition, older adults were observed during the visits to be less likely to engage in exercise and out-of-home activity (for example only 9 per cent of older adults were observed accessing out-of-home activities, compared with 72 per cent of adults with learning disabilities, significant at $p<0.001$).

As we discuss above, some of this might well be to do with preferences of OA, but the poorer outcomes in terms of occupation and social participation suggest that this is probably also about poorer quality of care, which has an impact on resident outcomes.

6.3 Nursing home quality scale and quality of environment

Those living in older adult homes experienced higher quality care, as measured by the nursing home quality scale (NHQS) ($p<0.001$). The mean total score for people living in OA homes was 116 with a range of 69-140. This compares with a mean score of 111 (range 69-132) for those living in homes for people with learning disabilities. Although this is significant, in reality the difference is small. On the subscales there was a slightly different pattern: on the Process subscale, OA homes had higher mean scores (88 for OA homes compared to 83 in LD homes ($p<0.001$)), whilst on the Structure subscale the learning disability homes had higher scores (mean 25 for LD homes and 23 for older adult homes ($p<0.001$)).

The scores on the nursing home quality scale can be categorised into low, good and high quality. Table 6.4 summarises the number of people in our sample living in homes in each category. Using these categories, there was a significant difference between residents of OA homes compared with LD homes ($p<0.001$), with those living in older adult homes more likely to be living in a home classified as being ‘good’ in terms of the nursing home quality scale, and those in LD homes more likely to be in a home classified as ‘low quality’.

Table 6.4: Nursing home quality scale and homelikeness scale

	Total sample N=173	OA only N=84	LD only N=89
NHQS (% of residents in homes rated as…)			
Low quality	22	13	30
Average quality	62	66	58
Good quality	16	21	12
Homelikeness scale			
Mean score[1]	2.47	2.66	2.29
(range)	(1 - 4.2)	(1 - 4.2)	(1 - 4.0)

1. A low score is more homelike.

In terms of homelikeness, those living in learning disability homes experienced significantly more homelike environments than those living in homes for older adults ($p<0.001$). This will be partly to do with size: the bigger the home the less homelike it can be on the scale used. There was a wide range of scores on this for both groups but most people were living in a home scoring three or below, which was at least quite homelike. Only 1.7 per cent of residents lived in a home that scored four or above.

Most people (70 per cent) lived in homes where the temperature was judged to be comfortable in all living spaces (bearing in mind the fact that people might be inactive). However, this meant that nearly a third of our sample were living in homes that were too hot or too cold for comfort. Three quarters of the residents were living in homes where all walking surfaces were even and not slippery, with most of the remainder in homes that were 'mostly' even and not slippery. Only 1 per cent were in homes where bedrooms or bathrooms were uneven or slippery. The problems, such as they were, tended to be in shared social spaces and hallways (5 and 4 per cent respectively). Nearly a third (31 per cent) of people lived in homes where they could not get out of the home unless accompanied and a further 25 per cent of people experienced some limitations in terms of access in and out of the home. This suggests that at least some of the lack of community involvement, participation and other out of home activities, reported above might be associated with limitations presented by the physical environment. However, only 1 per cent of people did not have access to a safe outdoor space.

6.4 Links with inspector quality ratings

We were interested in how our care quality indicators correlate with the quality ratings made by inspectors. Our Active Support Measure (ASM) correlated with percentage of lifestyle standards met or exceeded (r=0.302, p<.001). The Index of Community Involvement (ICI) and Index in Participation in Daily Life (IPDL) also correlated highly with the proportion of lifestyle standards met (r=0.750, p<.00,1 and r=0.657, p<.001 respectively).

The NHQS structure score correlated with percentage of standards met, both overall (r=.289, p<.001) and lifestyle standards (r=0.301, p<.001). Homelikeness also correlated with percentage of all standards met (r=-0.315, p<.001) and percentage of lifestyle standards met (r=-0.322, p<.001). This measure also correlated with the ratings of the KLORA Personal and health care support (r=-0.206, p=.007) and concerns and complaints (r=-0.206, p=.008).

6.5 Conclusions

The picture presented here about the quality of care homes and outcomes for residents as reflected in our measure of SCRQOL fits what we might expect to see. Needs in the most basic domains of people's lives are well met overall, but in areas such as social participation, occupation and control, residents are not as well supported as we might hope. When we link expected to current SCRQOL, we can see that care homes provide a high level of outcome overall, but the difference is much more limited in higher order domains. This applies to both types of home, but is more marked for OA than LD.

Our observational measures of care process and quality of care provide a helpful insight into what lies behind this. There is less contact, particularly less good active support from staff in OA homes than in LD homes. In part this may be to do with the lower costs of and associated lower levels of staffing in OA care homes.

At first sight it might seem surprising that the NHQS overall and process subscale identified higher quality in OA homes, given the evidence above that people in LD homes were receiving more active support. However, the instruments are measuring different aspects of care. The nursing home quality scale measures issues such as respect and dignity, and process issues such as dressing, together with environmental aspects such as homeliness, space, quality of environment and so on. The full name of the NHQS is the ‘Observable Indicators of nursing home care quality instrument’ - it is focusing on ‘care’ rather than one-to-one support.

Process in the NHQS is made up of items on respect, quality of interaction and dignity and support for personal care (eating, drinking, dressing etc). It is not surprising that these are better for older adults who will be less likely to have communication difficulties than those with LD. From other work (Beadle-Brown et al., 2009; Felce et al., 1987; Mansell ,2006) we know that people who have less severe LD have a better quality of life, more staff contact and so on. Staff find it more rewarding and easier to engage these people. Structure items are those related to environment, which is likely to be better in LD homes, which are smaller, more likely to have gardens and so on.

We found some associations between the measures of quality of care and support and the regulator quality ratings. In the next chapter we turn to the relationship between quality as measured by the regulator and the outcomes of care homes for residents.

7. Outcomes, quality and star ratings

In Chapter 5 we described how we can identify and monitor individual's capacity to benefit and the potential improvements care homes could produce: their capacity for benefit. Here we turn to examining how to measure and monitor what homes actually do produce, using the notion of quality in this context to reflect the level of outcome that is actually achieved. The aim is to find a way of relating what is achieved by homes, in terms of outcomes for residents, to current measures of quality or measures that could easily be instituted by commissioners, providers and/or the regulator. As we identified in Chapter 1, as the regulatory process was in the process of changing as the study was at the point of reporting, the aim was to investigate the relationship between outcomes and quality ratings to inform future developments rather than develop a cut and dried approach to linking outcomes to measures used at the time of the fieldwork.

In this chapter we start by clarifying how we are measuring outcomes for this purpose before describing the indicators of quality that are generated by the regulator and their relationship to one another. The results of a series of analyses are reported of the relationship between these measures and resident outcomes, and the role of other important contributing factors.

7.1 Measuring outcome

There are two aspects of outcome that are of interest:

- Current SCRQOL that residents are experiencing and
- SCRQOL gain: the difference between that and what they would have experienced in the absence of the care and support provided (expected SCRQOL).

In our analyses we explored the relationship between both of these and the regulator quality indicators, but our primary focus is on the latter.

Figures 7.1 and 7.2 show the distributions of the current SCRQOL experienced by OA and LD residents respectively. As we would expect the measure is highly skewed, but in the opposite direction to that of expected SCRQOL (see Chapter

5), reflecting, as we would hope, a high proportion of people with few if any unmet needs (see Chapter 6).

Figure 7.1: Current SCRQOL for OA

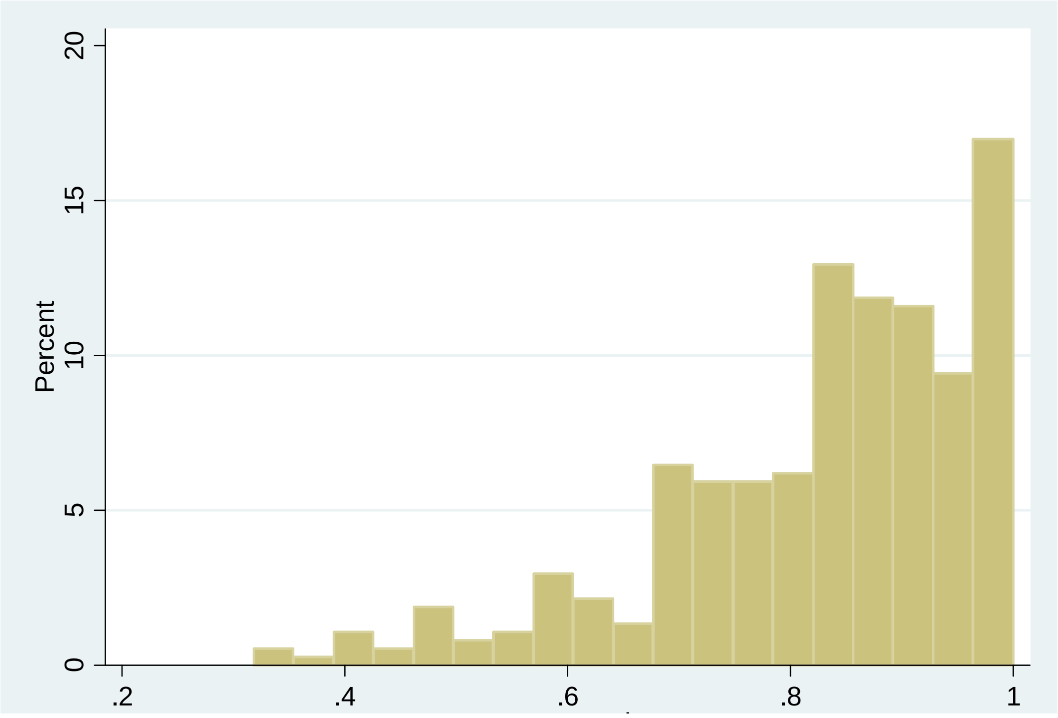


Figure 7.2: Current SCRQOL for people with LD

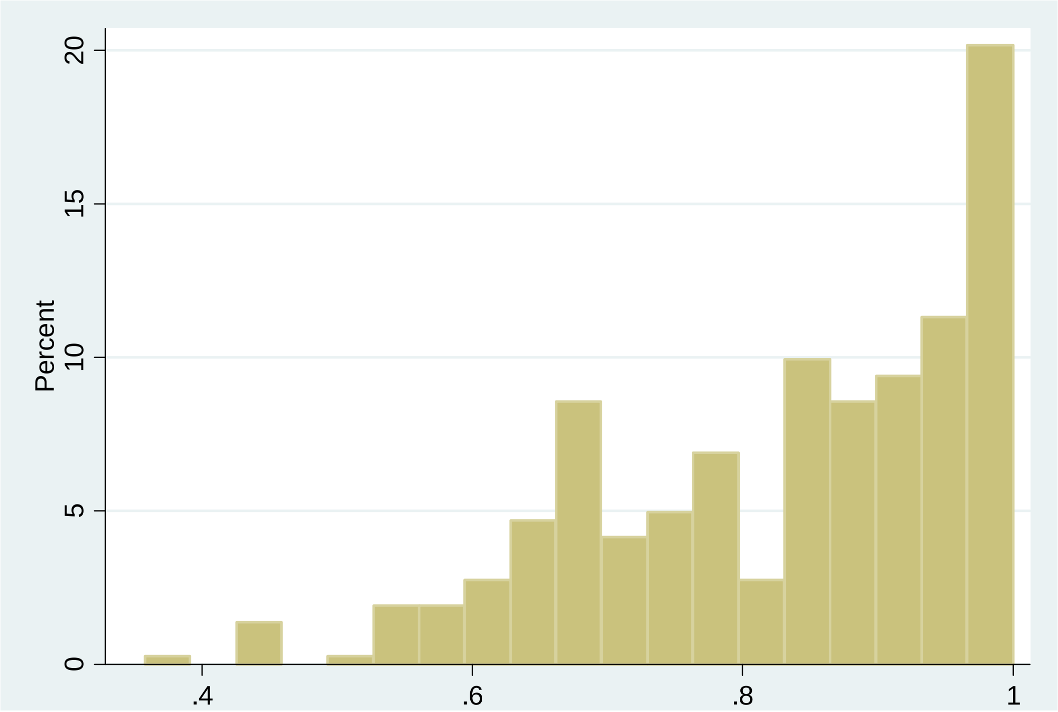


Table 7.1 shows the mean levels of current SCRQOL for each home type and by sector. People living in OA nursing homes had significantly lower current SCRQOL ($p<.001$). Taking the sample as a whole, there was a small but statistically significant higher level of current SCRQOL among residents of voluntary sector homes compared with private sector homes only ($p<.05$).

Table 7.1: Mean current SCRQOL by home type and sector

	Total sample	**Older adults**	**Learning disability**
Type of Home	**n =694**	**n=345**	**n=318**
Personal care	Mean = 0.85 (n=554)	Mean = 0.87 (n=219)	Mean = 0.84 (n=314)
Nursing	Mean = 0.77 (n=140)	Mean = 0.77 (n=126)	Mean = 0.83 (n=4)
Sector	**n=699**	**n=345**	**n=318**
Private	Mean = 0.82 (n=457)	Mean = 0.82 (n=258)	Mean = 0.82 (n=172)
Voluntary	Mean = 0.86 (n=183)	Mean = 0.87 (n=64)	Mean = 0.85 (n=112)
LA/NHS/Other	Mean = 0.83 (n=59)	Mean = 0.82 (n=23)	Mean = 0.86 (n=34)

Number of cases for LD and OA do not always equal total sample n because of missing information. Numbers shown are maximum n with necessary information for each group.

We constructed an outcome variable, SCRQOL gain, based on the preference-weighted outcome measure, which reflected the percentage of outcomes met: 0 representing the situation where no outcomes at all have been achieved and 100 where someone who would have had high level needs in all domains has all their needs met in all domains.[21] The distribution of this outcome for older adults (OA) and people with learning disability (LD) is shown in Figures 7.3 and 7.4 respectively. The mean level of outcome was 0. 576 for OA (median 0.586) and 0.555 for LD (median 0.561). In both cases, over 6 per cent of the sample achieved the maximum score of 1. Among OA, one person had a slightly negative score (minimum -.037): a situation where the care was actually undermining

---

[21] In the analyses we excluded the Dignity domain from the composite measures of current SCRQOL and outcome.

their well-being (see Chapter 6). Among people with LD, no-one had a negative score and less than 2 per cent had zero level of outcome.

There were no statistically significant differences on our SCRQOL gain measure of outcome between personal care homes and nursing homes, or any differences related to home sector.

Figure 7.3: SCRQOL gain for older adults

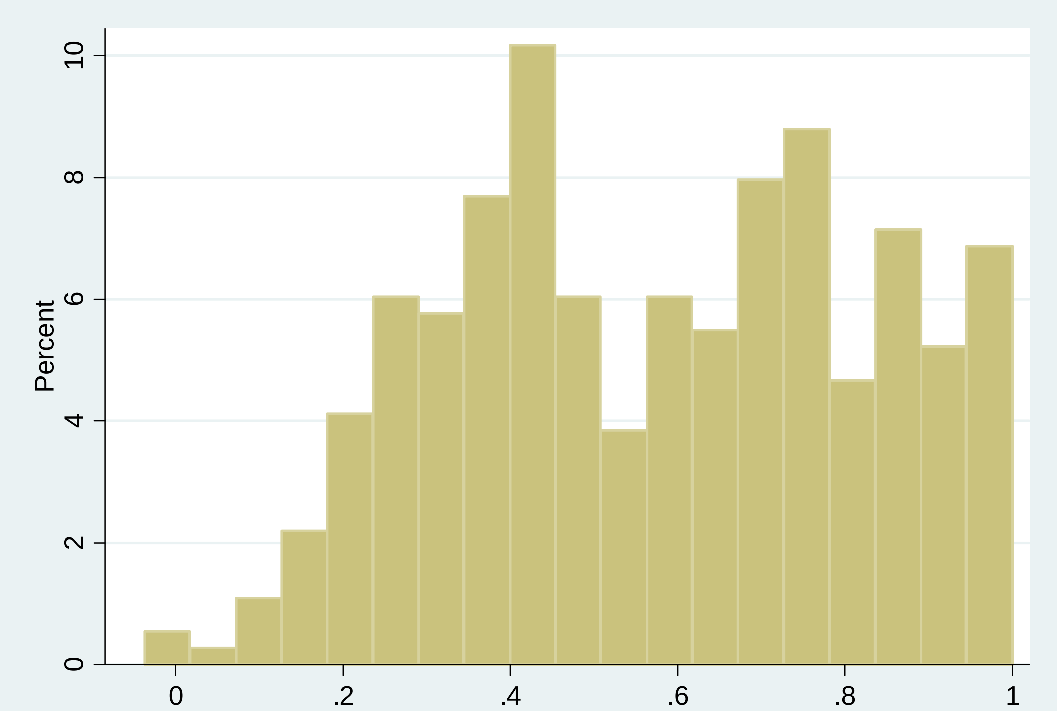
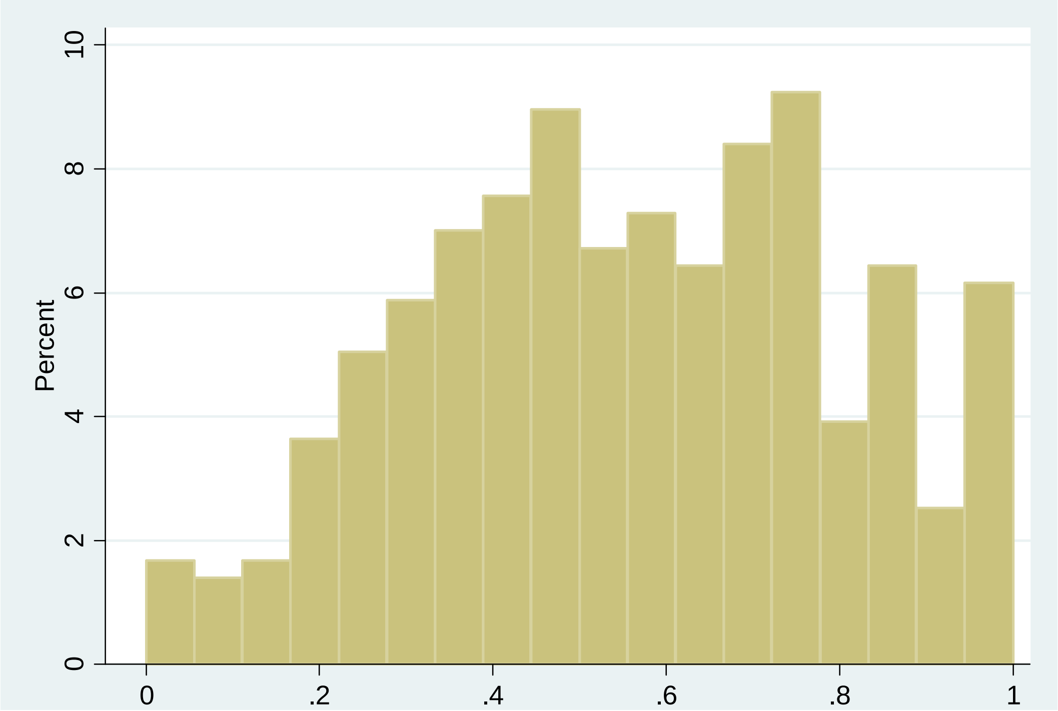


Figure 7.4: SCRQOL gain for residents of homes for people with learning disabilities



7.2 Quality ratings and standards

As we identify in Chapters 1 and 3, there are three interlinked regulator indicators of quality of care homes: care standards; what CQC term 'outcome' groups, and star ratings. There are 38 care standards for homes for OA and 43 for homes for younger adults, including LD homes. These standards are classified into eight 'outcome groups' whereby homes are inspected under the Key Lines of Regulatory Assessment (KLORA), each rated as 'excellent', 'good', adequate' or 'poor'. The distribution of the ratings of our sample homes across these outcome groups was described in Chapter 3.

The overall quality rating of the homes as zero (poor), one (adequate), two (good) or three-star (excellent) are, with a few exceptions,[22] based on an algorithm reflecting the ratings given for the outcome groups. Generally, as we would expect, the more 'outcome' groups that are graded as excellent, the more likely a service is to achieve three-stars. The more 'outcome' groups that are graded as poor, the more likely a service is to be rated as one or zero stars.

---

[22] For example, newly registered homes cannot receive an 'excellent' rating (a rating which means they are inspected only once every three years).

However, in awarding a quality rating, the regulator takes particular account of how safe and how well managed a care service is. As a result, services can only be as good as their 'poorest' rating in management and administration and complaints and protection.[23]

The history and relationship between these three types of quality indicator have important implications for the depth and quality of the data available. As we described in Chapter 3, virtually all the homes had star ratings and ratings of the ‘outcome’ groups. However, there were missing data on the more detailed standards and it was not entirely clear how recently these had been rated. In Chapter 2 we noted that we made every effort to get the ratings as contemporaneous as possible with the data collection in order to ensure links were as valid as possible. Thus, in the analyses described below, we focused particularly on the star ratings and the ‘outcome groups’ as the focus of inspection, resulting in more complete and reliable data. In the following we use the term KLORA rather than outcome groups to distinguish from our measures of outcome.

As a result of some early unexpected findings, for example, where good outcomes for residents were identified in zero-rated homes for OA, we examined the relationship between the KLORA groups reported and star ratings. We estimated an alternative rating for OA homes based on that predicted by the KLORA groupings. This resulted in 10 care homes for older adults being reclassified, usually (in seven cases) upgrading the homes by just one level[24]. We ran our models using both the adjusted and non-adjusted ratings. We report below the final model based on unadjusted ratings and report the impact of using the adjusted quality ratings.

7.3 Outcomes and quality ratings

In identifying the relationship between regulator quality ratings and outcomes for residents it is important that we allow for other expected causes of variation. We would expect individual personal characteristics to affect SCRQOL, with people

---

[23] www.cqc.org.uk/guidanceforprofessionals/socialcare/careproviders/inspection/ratingsandreports/qualityratings/ratingsrules.cfm.

[24] Initial exploratory analyses indicated that, in addition to the KLORA ratings, inspector characteristics, notably whether inspectors had a qualification in social work, mental health services or management, influenced the star ratings. Further analyses would be needed to investigate the impact of inspector characteristics on KLORA groupings and star ratings fully.

who are more highly dependent more likely to have a shortfall in current SCRQOL because of the extent or complexity of their needs, and we would expect higher levels of SCRQOL gain among those who have higher needs, and thus more capacity to benefit. We might also expect that the involvement and input of families would affect SCRQOL, although this is likely to be much more limited than among people living in private households. Age, gender and cultural factors associated with ethnicity[25] might be associated with SCRQOL, so we tested for these other demographic factors.

At the home level, in addition to regulator assessments of quality, we included the nursing home quality scale (NHQS) and subscales, size and sector of home, how long the service had been registered as a care home, whether the home was specialist and/or registered as a nursing home, and indicators of the general levels of dependency of residents (drawing on data about the wider sample of non-focal residents). We had intended to investigate whether there was a closer relationship when the observational approach to inspection (SOFI) was used but, as we identified in Chapter 3, this happened so rarely that it was not possible to investigate its effect.

Our underlying hypothesis was that variations in homes' quality affects outcome so, as we expected, initial diagnostic analyses identified that the intraclass correleation coefficient was highly significant suggesting that multi-level modelling was required. The general form of the model estimated was:

$$Y_{ij} = \beta_{00} + \sum_{g=1}^{t} \beta_g x_{gij} + \sum_{g=1}^{s} \gamma_g h_{gj} + e_{ij}\square + u_{0j}$$

Where:

Y_{ij} is SCRQOL gain of the *i*'th resident in the *j*'th home

β_{00} is the average intercept

x_{ij} are the characteristics of the *i*'th resident in the *j*'th home.

h_j are home level characteristics

e_{ij} reflects the random effect at resident level

---

[25] As described in Chapter 1, the issue of minority cultures was investigated in depth by one strand of the study, which is the subject of a PhD thesis and not reported here.

u_{0j} reflects the random effect at home level

Tobit (for current SCRQOL) and GLS (for SCRQOL gain, which was not seriously skewed, and for current SCRQOL) random intercept models were estimated using STATA 10. The results of Tobit and GLS models for current SCRQOL were very similar and we just report the GLS results here.

As we identified above, the star ratings and KLORA groups are all linked to the standards but the data on the standards themselves was incomplete in places. We focus here on the results of the analyses that used the star ratings as indicators of quality. In our analyses the few zero star homes were grouped with one-star homes.

Current SCRQOL and quality

We investigated the relationship between current SCRQOL, star ratings and KLORA groups as we hypothesised that we might get closer relationships than with outcomes as inspectors would more easily observe the current situation of resident, both directly and through interview, which would be expected to influence their ratings.

Among OA we found that individual residents' higher levels of dependency (ADL abilities) (p<.001), evidence of depression (p=0.039), communication problems (p=0.010) were all associated with poorer current SCRQOL. Once these were allowed for, a complex relationship between gender and age reached borderline statistical significance (male (p=0.080), age (p=0.104) interaction (p=0.053)). The relationship was such that, at the same age men had slightly better outcomes than women, but being older had a much more marked effect with older ages being associated with better current SCRQOL.

Using the star ratings as indicators of quality we found that, once individual-level characteristics were allowed for, two-star homes performed better than one-star homes but, surprisingly, three-star homes performed no better than one-star homes. We hypothesised that this might reflect the algorithm on which homes were classified. This was backed up by the model which used the KLORA groups to reflect quality. We found that good and excellent ratings on Personal and Healthcare KLORA were associated with better current SCRQOL (p=0.003 and p=0.019 respectively) and an excellent rating in management and administration associated with *poorer* current SCRQOL for residents, although this was of borderline significance (p=.059). As we describe above, the algorithm requires such a KLORA rating for a home to be rated as three-star, so it is not surprising that there is no association with homes being rated excellent and current SCRQOL for residents if there is an underlying association between poorer resident SCRQOL and highly rated management. We discuss this further in Chapter 8. In addition to the quality ratings by inspectors, the NHQS rating of quality of structure (p<0.001) and process (p=0.007) were associated with higher levels of resident current SCRQOL.

Among residents of homes for people with LD, the personal characteristics associated with poorer current SCRQOL were: needing help with activities of daily living (p=0.007), severe cognitive impairment (measured by the MDS CPS) (p=0.031), and the EQ5D item having problems with performing activities of choice (p=0.004). Getting any help from care home staff in keeping their accommodation clean and tidy was associated with better SCRQOL (p=0.042). Gender was not associated with SCRQOL but being younger was associated with higher levels of SCRQOL.

Again, the results in terms of star ratings were counter-intuitive, with two-star rated homes performing *worse* than one-star (p=.038) and three-star homes no better. In this model, higher quality reflected in the NHQS process indicator *was* associated with better SCQROL (p<.001), but including interactions between this and the star ratings did not result in any improvements in the model. The only KLORA group[26] that was associated with higher SCRQOL was, rather unexpectedly, choice of home – an area of quality that is about the level of information that people have about the home rather than reflecting current practice. It is possible that there is a better ‘fit’ between residents and the ethos of homes where there is good information available.

One possible interpretation of the counter-intuitive results of the relationship between current SCRQOL and star ratings, and the unexpected direction of effect of the Management KLORA group rating in OA homes, is that we are picking up the effect of high quality homes caring for people with higher CtB that we identified in Chapter 5. As we have identified above, we expect there to be poorer

---

[26] There were very few observations in some levels of quality ratings in the KLORA groups (see Chapter 3). Collapsing the categories for the purpose of the analysis may be one of the reasons no effect was found.

current SCRQOL for people who have higher CtB: that is, are more dependent, because it is more difficult to meet all of their needs. In order to allow for this, we turn to our preferred measure of outcome: SCRQOL gain.

SCRQOL gain and quality

Our preferred measure of outcome reflects the residents gain in social care related quality of life (SCRQOL gain): that is, the difference between residents' current experienced SCRQOL and expected SCRQOL in the absence of the help and support provided by the service. As with current SCRQOL we needed to allow for individual level factors and tested whether demographic characteristics and family input influenced outcome in addition to the expected effect of dependency factors. Tables 7.2 and 7.3 show the estimated models for OA and LD home residents respectively, when star ratings were used as the regulator indicator of quality (see Appendix F for the distributions of the variables included in these final models).

Table 7.2: Model of OA outcomes and home quality ratings[27]

	Coefficient	Probability
Resident characteristics		
Number of ADL problems (0 to 3)	5.04	<0.001
Ability to care for self (EQ5D-item: self care)		
(cf. no problems with self-care)		
Some problems with wash/dressing - All providers	16.67	<0.001*
Some problems with wash/dressing - Vol provider	6.55	0. 332*
Unable to wash/dress - all providers	18.02	<0.001*
Unable to wash/dress - vol providers	12.91	0.111*
Cognitive skills (MDS CPS item) (cf. no problems)		
Some problems		0.135
Moderately impaired	4.15	0.002
Severely impaired	9.46	<0.001
Home characteristics		
Quality (cf. zero/one star rated)		
Two-star rated home - all home types	4.80	0.077*
Two-star rated home - nursing home	-3.29	0.319*
Three-star rated home - all home types	10.14	0.001*
Three-star rated home - nursing home	1.38	0.861*
Number of places	-0.08	0.430
Type of home		
Voluntary sector	-0.16	0.954*
Nursing home	-3.60	0.187*
Constant	27.67	
Random effect		
Level 2 (care homes) variance	7.04	
Level 1 (residents) variance	16.04	

* Derived from bootstrapped standard errors; R^2 overall = 0.55, R^2 between = 0.71, R^2 within = 0.39; Wald chi square statistic =312.15, p<0.001

---

[27] Although coefficients not shown model included fieldworker dummies

Table 7.3: Model of LD outcomes and home quality ratings

	Coefficient	Probability
Resident characteristics		
Number of ADL problems (0 to 3) - all providers	2.67	0.055*
Number of ADL problems (0 to 3) – ind providers	4.16	0.002*
Ability to care for self (EQ5D: self care item) (cf. no problems)		
Some problems with wash/dressing	10.53	<0.001
Unable to wash/dress	15.13	<0.001
Cognitive impairment (MDS CPS score)		
Severely or very severely impaired (>4)	5.59	0.007
Communication (MDS CPS-item) (cf. understood)		
Rarely/never understood	10.44	<0.001
Home characteristics		
Quality (cf. zero/one star rated)		
Two-star rated home (mean home size)	-0.57	0.804*
Two-star rated home: adjustment for + 1 size (from mean)	-1.98	0.037
Three-star rated home (mean home size)	7.45	0.017*
Three-star rated home: adjustment for + 1 size (from mean)	-2.15	0.024
Number of places	-1.27	<0.001*
Type of home		
Independent sector	2.37	0.645
Constant	43.86	
Random effect		
Level 2 (care homes) variance	10.70	
Level 1 (residents) variance	14.45	

* Derived from bootstrapped standard errors; R^2 overall = 0.50, R^2 between = 0.67, R^2 within = 0.29; Wald chi square statistic =217.45, p<0.001

The model for OA explained 55 per cent of the variation in outcomes overall and 71 per cent of the variation in between homes[28]. As elsewhere in this report, we have shown the model that allowed for variations between fieldworkers. When fieldworker effects were excluded the overall R^2 was 0.37, indicating that 18 per cent of the variation was explained by variations in fieldworker judgements. For the LD sample, the model explained 50 per cent of outcome variation overall and 67 per cent of the variation between homes. In this case excluding fieldworker effects meant the overall R^2 fell to 36 per cent. The fieldworker effects were higher than in the CtB analysis reported in Chapter 5. This was to be expected, as the measure draws on both our measures of 'current' and 'expected in the absence of services' SCRQOL, and there was more variation between fieldworkers in judging 'current' or experienced' than 'expected' SCRQOL (see Chapter 4).

The level of outcome achieved from the use of care home services was found to be positively related to our measures of impairment for both OA and LD groups, including physical limitations (count of ADL problems and self- care difficulties) and cognitive impairment. These results are consistent with the hypothesis that services will deliver more outcomes for people with the most potential to improve their well-being.

For OA homes, the strength of the relationship between outcomes and impairment (in this case measured by ability to self-care) was found to be affected by the sector of the home. In voluntary sector homes, the people with self-care problems were not observed to have a significantly higher outcome improvement. This result is consistent with the results of the analysis of OAs capacity to benefit (CtB): namely, that voluntary homes seem to cater for, or attract, people where self-care related impairment appears less relevant (see Chapter 5).

Among LD homes, voluntary sector homes were not significantly differentiated from private homes in the relationship between impairment and outcomes. However, LA, NHS and 'other' homes, were different in this respect to private and voluntary sector homes. Here both private and voluntary sector home populations

---

[28] When we used the ratings adjusted to reflect the KLORA groups (see section 7.3 above) we were able to include 82 homes and 352 residents. The R^2 were virtually identical, with only overall R^2 slightly higher at 0.56.

showed a stronger relationship between ADL impairment and outcome scores than the other types of home.

Although different home types appear to cater for different groups of people, in both OA and LD homes, sector on its own was not significant, meaning that outcomes for any person in a voluntary sector home were no better or worse than people in other home types.

Turning to the relationship between SCRQOL gain and quality rating scores, for LD homes, simple univariate comparisons showed that three-star homes had significantly higher mean SCRQOL gain scores than zero and one-star homes, although this was only at the 10 per cent significance level. Two-star homes also had a slightly higher, but not significantly different, mean. For OA homes, the relationship in univariate analysis was different. Two-star homes showed significantly better outcomes than zero or one-star homes but, as in the current SCRQOL model, three-star homes were not significantly better than these poor or adequate homes.

Nonetheless, when allowance was made for client characteristics and home size in multivariate analysis, three-star homes did deliver significantly better outcomes, other things being equal, for the OA sample and the LD sample, than two-star (and zero/one-star) homes. Three-star homes showed significantly better outcomes than zero or one-star homes in both models. Two-star homes also showed significantly higher outcomes than zero and one-star homes for OA, but not for LD homes.

For OA homes, the relationship between star rating and outcomes appeared to be dependent on the registration type of home. The significance of the relationship between star ratings and outcomes was driven by homes registered for personal care only. Nursing homes did not show significant relationships between star rating and outcomes. There was also some evidence that nursing homes delivered less SCRQOL gain, once factors such as resident dependency were allowed for, when we used the star ratings adjusted to reflect the KLORA groups (see above) in the model ($p=0.066$).

For LD homes, the relationship between star rating and outcomes appeared to be dependent on the size of the home. For large homes (that is larger than the mean size) the association between star rating and outcome was lower. Table 7.3 shows an adjustment to the size of the effect of star rating on outcomes for larger homes. For example, for the mean size a three-star home had a 7.5 per cent higher outcome score than a 0/1 star home. But for a three-star home which is one or more place larger, the size of this relationship fell by 2.15 per cent.

Although star ratings did show significant effects for homes, the size of the effects were small, improving overall explained variation by only a couple of percentage points (for both OA and LD), accounting for about 5 per cent of the explained variance.

When we explored KLORA groups as indicators of quality, the health and personal care group was associated with outcomes for OA home residents (p=0.009), but only at a borderline level of statistical significance (p=0.055) when fieldworker effects were included. The length of time that the home had been registered and size of home were also statistically significant in this model (p=0.021 and p=0.033 respectively) but both were non-significant when fieldworker effects were included. When we investigated the KLORA groups in homes for people with LD, we found one association: higher outcomes were associated with an excellent rating on Individual needs and choices (p=0.023). The significance of this effect disappeared once fieldworker effects were included.

7.4 Fees and outcomes

As we described in Chapter 1, an overarching aim of the work is to provide a basis for linking outcomes to either extant or easily collected data so outputs can be weighted to reflect changes in the value of services. An important reason for doing this is to provide a basis for monitoring changes in productivity: are we getting better or worse value for money over time? It is of interest to investigate whether costs, which in the case of care homes are fees to those who purchase the services, are related to our measures of value.

We undertook an initial investigation into the relationship between fees and our measures of capacity to benefit and outcome, measured through SCRQOL gain. Table 7.4 shows the resulting model for OA outcome. As we would hope, when we exclude quality ratings, higher fees are associated with higher outcomes. Average fees were about £500 so the effect size was not very large - an additional £50 per week would purchase one percentage point additional outcome, for example. There was a similar association with CtB (p=0.044) with higher prices associated with increased capacity to benefit, over and above individual level dependency indicators.29

A similar analysis of price effects for LD homes also found associations but very high price homes were associated with worse outcomes. It is likely that these specialize in caring for people with very challenging behaviour. Overall, although the net association between price and outcome was positive, it was not significant (Table 7.5).

---

[29] The same dependency indicators are shown in Table 7.4

Table 7.4 OA homes' model of fees and outcomes

	Coefficient	**Probability**
Resident characteristics		
Number of ADL problems (0 to 3)	3.87	0.003
Needs help dressing	10.48	0.002
Needs help with personal affairs	5.29	0.085
Cognitive skills (MDS CPS item)		
Some problems	4.31	0.154
Moderately impaired	9.64	0.003
Severely impaired	11.32	0.002
Price		
Weekly fee	0.02	0.041
Home characteristics		
Number of places (squared)	-0.0007	0.316
Type of home		
Voluntary sector	-1.89	0.589
Nursing home	-5.98	0.063
Constant	30.63	
Random effect		
Level 2 (care homes) variance	7.29	
Level 1 (residents) variance	16.70	

R^2 overall = 0.52, R^2 between = 0.79, R^2 within = 0.36; Wald chi square statistic =278.97, $p<0.001$; n = 334

Table 7.5: LD homes' model of fees and outcomes.

	Coefficient	Probability
Resident characteristics		
Number of ADL problems (0 to 3) - all providers	4.76	<0.001
Problems with mobility	5.29	0.013
Cognitive impairment (MDS CPS score)		
Severely or very severely impaired (>4)	7.15	0.001
Problem behaviour: frequently	2.35	0.368
Communication (MDS CPS-item)		
Rarely/never understood	11.90	<0.001
Price		
Weekly fee	0.02	0.331
Weekly fee squared	-6.59E-06	0.331
Home characteristics		
Number of places	-0.91	0.034
Type of home		
Independent sector	-4.57	0.467
Constant	42.03	
Random effect		
Level 2 (care homes) variance	11.46	
Level 1 (residents) variance	14.51	

R^2 overall = 0.26, R^2 between = 0.58, R^2 within = 0.45; Wald chi square statistic =170.04, $p<0.001$; n = 307, care homes n = 87

7.5 Conclusions

What are the implications for these findings for using information collected by the regulator for monitoring quality, defined as the degree to which outcomes are achieved by residents? Clearly there is a relationship between quality of care provided and outcomes, and this is reflected in the relationship between star ratings and outcomes for residents. However, the basis of the star ratings currently puts a great deal of emphasis on sound management and administration. This means there is a lot of 'noise' in the relationship and that any indicator based on these star ratings would need to take this into account. Moreover, as we identify above, the relationship between quality as measured through star ratings and current SCRQOL is likely to be influenced by the finding from Chapter 5: that higher quality homes care for people who have higher CtB in ways not fully captured by our individual measures of abilities in activities of daily living and so on.

While it is disappointing in terms of developing an indicator that can reflect quality across all types of home, the lack of relationship between quality ratings of OA nursing homes and suggestions of poorer SCRQOL outcomes for residents of nursing homes may well reflect an emphasis both in the homes and by the regulator on health rather than social care factors and quality of life. In some ways, the relationship with the KLORA groups are also rather disappointing : we do not see the strength of relationship and explanatory power that we might hope. Moreover, the negative relationship between the Management and administration grouping and OA resident outcomes is a little worrying in terms of using this as an indicator: we would not want to create perverse incentives in any system set up. Nevertheless, it is worth considering how regulator data including quality ratings could be used to monitor the value of what is provided and feed into calculations of productivity. The initial analysis of the relationship between our indicators of the value of services were encouraging, at least for OA. In the next chapter we explore the implications of the findings for developing our tool for measuring and monitoring outcomes: ASCOT.

8. Conclusions and looking forward

The aim of the study was to develop an approach to measuring outcomes for care home residents and demonstrate how these could be linked to easily collected, or already extant, routine sources of data. The study was undertaken at a time of change for the regulator, with the formation of the joint health and social care regulator, the Care Quality Commission (CQC) and a consultation exercise is being conducted on the regulatory process, including plans to drop the care standards that form the basis for the current KLORA groups and star ratings. The project was designed to make links with the information currently available through the regulatory process. While, as we describe below, this means the study provides a valuable potential source of information to the regulator, the planned changes mean that there is little to be gained through detailed analyses of exactly how the relationships with current regulatory data could be used in the future to monitor outcomes over time. In terms of examining the implications for the future of what we have learnt, we therefore concentrate on four main aspects: the tool that has been developed; the monitoring of CfB and quality using data that are or could be available through the regulator; the potential use for commissioners and the potential use for National Accounts. In our discussion we also identify the implications of other developments both in ongoing research and national data collections. We start by identifying the study's limitations in order that these can be kept in mind in considering the implications of our findings.

8.1 Limitations of the study

Inevitably with a project of this ambition, size, and complexity, there were a number of limitations. While every effort was made to limit the impact on the participating homes, the prospect of a two-day observation-based fieldwork was a deterrent to many homes. While we achieved our aim of reflecting the range of home type in terms of quality, sector and registration, the response rate was low. However, the overall quality distribution was reasonably representative nationally: the only excluded group being poor rated homes for people with a learning disability, which represent less than 2 per cent of homes nationally.

Once the homes were recruited, the response rate among residents was good, and for the most part the focal resident sample appeared to be representative of the wider home population and what we would expect from other studies. However, there was some concern that the sample in OA nursing homes was biased towards the more able residents. While they matched a contemporaneous sample in another study in terms of EQ5D scores, this study might also have been picking up a more able sample. Other evidence suggested that the AQAA data supplied by the home to the regulator, from which we sampled residents, may have been drawing on the more able residents.

While analysis of our measure suggested good internal reliability, inter-rater reliability was not as good as we might have wanted. However, it was interesting to note that in both instances reliability was higher for what might be seen as the more difficult to establish indicator of expected SCRQOL in the absence of services. Our analyses allowed for variations between raters and, as we discuss below, evidence from this study could feed into improving this in future studies.

8.2 The research toolkit and measure

The Adult Social Care Outcome Toolkit (ASCOT) will include a variety of instruments and methods for establishing social care related quality of life outcomes in a variety of contexts. The research tool for application in a care home will draw on the research reported here and on the results of other research projects developing the ASCOT measure (Caeils, 2010; Burge et al., 2010). We summarise what we would include from the toolkit used in this study to generate the ratings for each domain, drawing on evidence from the fieldwork process and the analyses of the data.

While there was clearly room for improvement in terms of inter-rater reliability (see Chapter 4), for the most part the analysis suggested that judgements were being made on a reliable basis with fieldworkers building up expertise over the course of the project and providing valuable evidence to justify their ratings (see Appendix E). This evidence drew on what residents themselves said, when they were able to communicate, and staff views and observations made during the visit.

The toolkit will include an interview schedule for residents. It is important to ask residents themselves about their perceptions of their current levels of social care-related quality of life (SCRQOL) and, if all residents were able to respond to such questions, in principle we would take these responses as the ‘gold standard’ measure of current status. We acknowledge the potential for adaptation and thus for people to overstate their current quality of life, but individuals’ own perspective has to be the starting point both ethically and from the point of view of creating the right incentives in collecting and interpreting data. However, as we show in Appendix F, a high proportion of residents were unable to take part in an interview at all, and the nature of residential care makes it particularly difficult to ask residents about their ‘expected’ situation in the absence of the service.

Interviews with key staff would also be important as they will know residents well and, in addition to identifying what they feel residents current SCRQOL and expected SCRQOL to be, are able to provide evidence to back those views up. This helps in the interpretation of both current observed behaviour and in deciding whether the basis for those views is valid given our definitions. The staff interviews could also include other items for reflecting needs, such as activities of daily living (ADLs), EQ5D and the Minimum Data Set Cognitive Performance Scale (MDS CPS).

While some summary aspects of the Nursing Home Quality Scale (NHQS) that we used worked well and might prove helpful to include, others were less so. We would not include detailed evidence on ‘homelikeness’, as this can result in people focusing on those aspects of homes that are more likely to affect visitors’ views than residents’ experiences. However, evidence about the nature of the interactions with staff is important, and we would seek to ensure that these elements were incorporated in the toolkit.

From feedback from the fieldworkers and analysis of the basis of the evidence used, it was clear that a period of detailed observation was essential to making judgements about people’s current SCRQOL and expected levels in the absence of the service. It is easy to be over-influenced by the ‘halo effect’ of a generally well-run home and a pleasant environment, and miss what individuals’ actual experiences are. As we identified in Chapter 2, the period between 4pm and 6pm covers a period that usually offers the most opportunities for staff to interact and support the residents thus maximising the opportunities to observe engagement and the quality of staff support but at the same time we know that in terms of observed engagement this period is consistent of other times during the day*.* The Engagement in Meaningful Activities and Relationships (EMACR) measure used in this study worked well in identifying what people did and the quality of interactions. However, the classifications used in summarising the results of this observational period – the Active Support Measure (ASM) - were in some areas of evidence more technical than is required for identifying SCRQOL and in others provide evidence for identifying SCRQOL. In essence the ASM is a measure of the quality of staff support for engagement in meaningful activities and relationships and person-centred support. It is a process measure related to quality of service rather than a measure of user outcome per se. However, forcing fieldworkers to think about the quality of staff support did help them to think about whether people's needs were being met. We would revisit this and adapt it to create a new 'Support Quality and User Experience' rating. The recommended codings would include items more closely related to SCRQOL (such as loneliness), evidence of quality of care, drawing on individualised aspects of the NHQS, and evidence of distress, happiness and pain.

Training in the observational aspects and identifying current and expected SCRQOL from the accumulated evidence will be essential. The training would draw on the evidence provided by the fieldworkers, particularly by the two who coded in a systematic and consistent way and other lessons learnt from the research, including the emphasis on same-day ratings, need for good evidence and focus on the individuals. The training process itself could also benefit from lessons learnt about careful matching of 'buddies' and the potential impact of poor performers: the poor performers often brought down their buddy's performance too.

The Dignity domain, reflecting the impact of the process of care, was not included in our final analyses as we were not entirely confident of the evidence provided (see Chapter 4) and as it was challenging to include as, by definition, it was not applicable 'in the absence of services'. Nevertheless, we would aim to include this in the toolkit as an important aspect of service users SCRQOL, drawing on the results of this study and other work to refine the approach. In more recent versions of the measure (which have four levels in each domain) we have included a level where services are identified as having no impact on recipients' self esteem (Netten et al.,2009). The weighting for this 'no impact' level can be incorporated in the total 'expected' SCRQOL score. The outcome is then the difference between this and the overall experienced SCRQOL score, which would include the currently experienced Dignity domain.

Another piece of work, which was developing the ASCOT outcome measure further (the Outcomes of Social Care for Adults (OSCA) project), identified the need for an additional response option, over and above the three basic no, low and high needs options used in this project, to reflect a higher level of SCRQOL in each domain (Netten et al., 2009). Social care policy and interventions are increasingly focused on the wider 'well-being' agenda (Department of Health, 2008) so it is important that any measure of social care outcomes should be able to reflect the improvement in people's *capability* as well as basic levels of *functioning* (Sen, 1985). Thus for example, rather than just identifying whether people are sufficiently occupied, we would hope to distinguish whether people were spending their time doing want they wanted to do. This distinction between a high level of SCRQOL and basic no needs 'mustn't grumble' level will add a further challenge, but is one we would aim to encompass in the revised toolkit.

Preference weights generated from previous work (Burge et al., 2006) were used for the purposes of analysis in this report. The toolkit will incorporate preference weights from the preference study being undertaken as part of the MOPSU project (Burge et al., 2010).

8.3 Monitoring capacity for benefit

We identified in Chapter 5 that service-level capacity for benefit (CfB) is based on individual resident capacity to benefit (CtB) from that service, and reported on the best models for predicting this. We focus here on the individual-level characteristics and discuss in section 8.4 the implications of the home-level characteristics influence on CfB when discussing the monitoring of quality.

If we want to monitor the impact of changes in the care home population on CfB of homes in the future, it will be important to consider:

- Whether information is to be collected at an individual or home population level (e.g. the number/proportion of residents who are bedfast)
- Who does the data collection and what this implies for incentives and potential gaming
- The level of accuracy required
- The resource implications

Clearly there will be trade-offs: the more accurate and independent the approach, the more resource-intensive it is likely to be.

If information was to be collected at the individual level then the resident characteristics included in our models could be used to predict CtB and the average of resident CtB. For smaller homes this might be for all residents and for a random sample of residents in larger homes. The data collected could be just those variables we have identified (about 10 items) and could be the same for all care homes for adults:

- EQ5D (anxiety/depression, mobility and self-care items)
- MDS CPS (Communication question)
- ADLs, including at least:
    - Wash face and hands
    - Get in and out of bed (or chair)
    - Getting dressed
- IADLs
    - Needs help with personal affairs

We might also consider collecting data about contact or input from families, which might prove useful in the future, even if our current models did not include this variable.
Alternatively, more or slightly different sets of items, which reflect the core areas identified in the model, but are seen to be more widely useful, might be collected. There might be a wider interest in monitoring EQ5D, a widely-used health outcome measure, as a whole. Another alternative is to go for a ‘minimum data-set’ by reducing the level of overlap in areas such as washing and dressing so keeping the number of items as small as possible. We would argue that it is likely to be more useful to be consistent across different user groups, but the data collected could be restricted further if it were made user group-specific.

We identify above that our OA nursing home resident sample is significantly more dependent than personal care home residents but that it is possible that our sample is more able than the national average. Our sample included the full range of dependency levels, and the measures we propose here to monitor resident CtB are able to reflect the full range of dependency, so we would be confident that they could reflect changing CfB of the care home population.

Ideally future work would validate the relationship with a fully representative sample, or at least a more representative sample since ‘fully representative’ is probably an even less achievable gold standard than in most other settings.

As we described in Chapter 3, the regulator collects, through the Annual Quality Assurance Assessment (AQAA) form, the number of people in each home who are bedfast, require help with dressing, washing/bathing, going to the toilet, are incontinent, have dementia and so on. Potentially this would be a source of information about the resident population that could be used to monitor the population of care homes, and this might be seen as less resource intensive than collecting data at an individual resident level. However, the concerns we have raised about the reliability of these data suggest that careful consideration would be needed before placing much faith in data collected in this way.

We have not pursued further analysis of the resident sample here, but, should the regulator wish to pursue collecting information at an individual level, this provides us with the potential to investigate the implications for sample sizes for larger homes. In developing an approach to monitoring CfB of homes over time, we could draw on our models to identify information that could be collected at an individual or home population level.

Our assumption throughout the analysis has been that everyone - regardless of their level of needs - could, through the support of care home services, have all their needs met in each domain. This might be very costly and challenging for some but potentially the service would want to deliver this for every resident: the quality measure would reflect how far this aim was actually achieved in practice. However, it is less clear that homes could be assumed to be aiming to deliver higher level outcomes in all domains – what we have described as what people want ideally rather than the ‘mustn’t grumble’ in other developmental work (Netten et al., 2009). We will need to consider how we define the top level of capacity for benefit of services for the future. This could be something that homes themselves specify, that we identify through observation, or that we assume for all services of a given type.

8.4 Monitoring quality

We have defined quality as the degree to which people are achieving outcomes: the difference between experienced SCRQOL and expected SCRQOL in the absence of services. Quality is more often defined in terms of structure and process, and this is reflected in the care standards and KLORA ‘outcome’ groups used by the regulator. While we would anticipate that better structure and processes should be associated with better outcomes, it is often argued that the former are necessary but not sufficient factors so it is perhaps not surprising that we did not find very strong relationships between the regulator quality indicators and our outcome measures.

The cobweb charts in Chapter 6 demonstrate that homes deliver substantial outcomes in terms of SCRQOL gain, as we would hope, given the levels of impairment and vulnerability of the resident population that they serve. However, while more basic aspects of SCRQOL are met, there are clearly unmet needs in higher-order aspects of SCRQOL, both in homes for OA and people with LD. The reasons for this became clear when we examined the observational data on what was happening to people in the homes. While the quality of the interactions in terms of respect and friendliness were better in OA homes, actual levels of activity and active support, to maintain activity were particularly low in these homes. While there was more evidence of active support, the situation was not much better in LD homes.

Individual characteristics that influenced SCRQOL were primarily related to levels of dependency, although there were some associations between age and currently experienced SCRQOL. However, our primary interest was in the association between home-level characteristics, particularly quality ratings, and outcomes.

An important complicating factor in this relationship was the unexpected relationship between quality of home and residents’ CtB . The analysis in Chapter 5 identified higher CtB among residents of higher quality homes once individual dependency-related characteristics had been allowed for. Our hypothesis is that this reflects the additional impact of residents for whom it is difficult to provide care and support. Higher quality homes are better able to cope with these residents than poorer quality homes.

This has a knock-on impact on our analysis of quality ratings and outcomes. Current experienced SCRQOL is lower among those with higher CtB, not surprisingly, as it is more difficult to meet all the needs of those with higher level needs. While individual dependency characteristics pick up some of this effect, we might expect the quality ratings to reflect the impact of caring for these more ‘difficult’ residents. This works in the opposite direction to the expected impact of higher quality of provision, resulting in inconsistent results. Another potential contributing factor to the inconsistent results is the emphasis currently given to good management practices in the star ratings. The analysis of the relationship between current SCRQOL and KLORA groups suggested that ‘excellent’ management was associated with poorer SCRQOL among residents of OA homes, and it is possible for a home to be rated as three-star with an excellent rating in this KLORA group and ‘good’ ratings in all the other KLORA groups. While it is possible that this reflects an emphasis on administrative process rather than resident outcomes, we need to be careful in our interpretation given the association between CtB and quality described above.

The inconsistency in the relationship between quality ratings and resident outcomes is not evident when we turn to SCRQOL gain, where those with higher needs are identified as having higher outcomes. We have a consistent picture in terms of star ratings, although this does not explain a great deal of outcome variation and was not evident for OA nursing homes. There was also some evidence that they tended to deliver lower outcomes once allowance had been made for levels of dependency. These last results suggest that possibly both nursing homes and the inspections of those homes are more focused on health than social care, something that the regulator might want to consider in reviewing the regulatory process.

Voluntary homes did not deliver significantly different outcomes than other sectors. However our CtB analysis suggested that these homes were caring for people less dependent on the service for their outcomes, and this was demonstrated for both specialist homes for people with LD and homes for OA. Thus we may not be comparing like with like. We hypothesised that the different levels of dependency might reflect greater trust in such homes, so that people are more willing to move into them at lower levels of dependency. It is also possible that these homes can draw on other inputs and support beyond the direct service delivered.

This complicated picture does not readily lend itself to identifying a way of monitoring quality in terms of achieving outcomes. As we identify above, the regulator is reviewing the regulatory process so at this stage it is more appropriate for the results to feed into that than to devise an approach based on regulatory data that will not be available on the same basis in the future.

The ASCOT measure is currently being tested as an important component of the annual social care services user experience survey (UES), which from 2011 is planned to cover all services, including care homes. This takes the form of a self-completion questionnaire which is being tested for the care home context. Potentially this is a source of information on current SCRQOL that could be used to reflect quality of provision, although it would be important to draw on the lessons from this study in the interpretation of any results.

8.5 Commissioning

While there is clearly a relationship between the regulator's quality ratings and outcomes, lack of much explanatory power for individual- level outcomes is a drawback when it comes to commissioning. Nevertheless, it would be feasible to use information about the characteristics of individuals (ADLs, whether they have communication difficulties, cognitive impairment levels and so on, as reported in our models) in commissioning, to purchase intended outcomes. In terms of quality (the degree to which the outcomes are delivered), commissioners could either pay rates to homes related to their star ratings to reflect that outcomes are related to these ratings or conduct surveys of residents to identify current levels of SCRQOL. Our analysis of the relationship between fees and CtB and outcomes suggested that this may be happening to some degree already.

As we identify above, cognitive impairment and communication difficulties necessarily mean that many individuals will not be able to respond to questionnaires or interviews. Nevertheless, help from advocates, relatives or others who know them well could help provide information about residents' current SCRQOL. The proposed UES might be used as a general source of information to monitor how outcomes are being delivered to service user groups, be increased to get an indicator at home level or form part of reviews or other contract-related follow-ups at an individual level. The toolkit will include interview and self-completion versions that could be incorporated into such processes.

8.6 National Accounts

The key to any inclusion in National Accounts would be consistency and availability of data over time. In light of the current review underway, it is not sensible to develop a detailed approach for reflecting quality-weighted outputs based on current data provided by the regulator. However, the results suggest that the basic approach of collecting information about residents and rating quality provides the potential for this to be put into practice. Information about the nature of the care home population could be used to reflect changes in capacity for benefit, and are likely to be of wider value for policy in terms of monitoring the use of this expensive resource. In terms of quality, output could be weighted to reflect the proportion of residents in homes with higher star ratings. Further adjustments might be done to reflect home characteristics, such as size. The actual degree of explanatory power is currently low, but the analysis and data from this study could potentially feed into making future data reflect the relationship with CfB and outcomes more closely than is currently the case. Alternatively, the proposed Putting People First (PPF) UES could be used at a national level as a regular source of data on current SCRQOL experienced in care homes.

The quality weighting resulting from these data would never be perfect, but there might well be a point at which indicators that are available of CfB and quality are seen as a better indicator of changes in output than fees, both conceptually and in allowing analysis of changes in productivity. The distinction between CfB and quality facilitating an understanding of whether observed changes are about changes in residents served or outcomes delivered.

8.7 Conclusion

In devising an approach to measuring the outcomes of care home residents, the study has addressed a challenging and important topic. While inevitably there are some reservations, the analyses have identified important relationships and the potential for regulatory data to be used to monitor the changing value of the outputs of care homes over time.

9. References

Aman, M. G., W. H. Burrow, W. H., and Wolford. (1995) The Aberrant Behavior Checklist-Community: Factor validity and effect of subject variables for adults in group homes *American Journal on Mental Retardation* 100(3), 293-292.

Aud M., Rantz, M.J., Zwygart-Stauffacher M., Flesner M. (2006) Measuring quality of care in assisted living: A new tool for providers, consumers, and researchers. *Journal of Nursing Care Quality* 22(1),4-7.

Beadle-Brown, J., Hutchinson, A. and Mansell, J. (2008a) Care standards in homes for people with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities,* 21, 210–218.

Beadle-Brown, J., Hutchison, A and Whelton, B. (2008b) A better life: The implementation and effect of person-centred active support in the Avenues Trust, *Tizard Learning Disability Review*, 13 (4), 15-24.

Beadle-Brown, J., Murphy, G. and DiTerlizzi, M. (2009) Quality of Life for the Camberwell Cohort: A Brief Report, *Journal of Applied Research in Intellectual Disability*, 22, 380-390.

Bowling, A. (1995) Measuring Disease, Open University Press, Buckingham.

Burge, P., Gallo, F.and Netten, A. (2006) *Financial valuation of PSS outputs and quality changes*, PSSRU Discussion Paper No. 2338, Personal Social Services Research Unit, University of Kent.

Burge, P., Potoglou, D., Kim, C., and Hess, S. (2010) *How do the public value different outcomes of social care? Estimation of preference weights for ASCOT*, RAND working paper, Rand Europe, Cambridge.

Caiels, J., Forder, J., Malley, J., Netten, A. and Windle, K. (2009) *Measuring the outcome of low-level services: Final Report*, PSSRU Discussion Paper, Personal Social Services Research Unit, University of Kent.

Care Quality Commission (2010) *Consultation: Assessments of quality in 2010/11*, Care quality Commission, London.

Cheung, K., Oemar, M., Oppe, M. and Rabin, R. (2009) *EQ-5D, Basic Information on how to use EQ-5D*, EuroQol Group.

Commission for Social Care Inspection (2005) *Inspecting for Better Lives – Delivering Change*, Commission for Social Care Inspection, London.

Commission for Social Care Inspection (2009) *The State Of Social Care In England 2007-2008*, CSCI Report, Commission for Social Care Inspection, London

Conroy, J. W. and Feinstein, C. S. (1986) *The Choice-Making Scale*, Conroy and Feinstein Associates, Philadelphia.

Curtis, L. (2008) *Unit Costs of Health & Social Care 2008*, Personal Social Services Research Unit, University of Kent.

Darton, R. (1986) *Methodological report of the PSSRU Survey of Residential Accommodation for the Elderly, 1981* Discussion Paper 422, Personal Social Services Research Unit, University of Kent at Canterbury.

Darton, R., Miles, K. and Netten, A. (1998) *Cross-sectional survey of residential and nursing homes for elderly people. Description of the database*, Discussion Paper 1391, Personal Social Services Research Unit, University of Kent at Canterbury.

Darton, R., Netten, A. and Brown, P. (1997) *A longitudinal study of admissions to residential and nursing home care following the community care reforms*, Discussion Paper 1417, Personal Social Services Research Unit, University of Kent at Canterbury.

Darton, R., Forder, J., Bebbington, A., Netten, A., Towers, A.-M. and Williams, J. (2006) *Analysis to Support the Development of the Relative Needs Formula for Older People: Final Report*, PSSRU Discussion Paper 2265/3, Personal Social Services Research Unit, University of Kent, Canterbury.

Darton, R., Forder, J., Netten, A., Bebbington, A., Holder (née Williams), J., Towers (née Muncer), A-M., (Forthcoming) Slicing up the pie: allocation of central government funding of care of older people, *Social Policy & Administration.*

Department of Health (2008) *Putting People First*, Department of Health, London.

De Wit, G.A., Busschbach, J.J.V. and De Charro, F.T. (2000) Sensitivity and perspective in the valuation of health status: whose values count? *Health Economics*, 9, 2, 109-126.

Felce, D., Saxby, H., de Kock, U., Repp, A., Ager, A. and Blunden, R. (1987) To what behaviors do attending adults respond?: A replication. *American Journal of Mental Deficiency*, 91, 496-504.

Feldt, K. S. (2000) Checklist of Nonverbal Pain Indicators. *Pain Management Nursing*, 1(1), 13-21.

Forder, J., Towers, A-M., Caiels, J., Beadle-Brown, J. and Netten, A. (2007) *Measuring Outcome in Social Care: Second Interim Report*, PSSRU Discussion Paper No. 2422, Personal Social Services Research Unit, University of Kent.

Glendinning, C., Challis, D., Knapp, M., Manthorpe, J., Netten, A., Wilberforce, M., Fernandez, J-L., Jacobs, S., Jones, K., Moran, M and Stevens, M (2008) *Evaluation of the Individual Budget Pilot Projects: Final Report.* Social Policy Research Unit, University of York, York.

Hatton, C., Emerson, E., Robertson, J., Gregory, N., Kessissoglou, S., Perry, J., Felce, D., Lowe, K., Walsh, P. N., Linehan, C. and Hillery, J. (2001) The adaptive behavior scale-residential and community (part I): towards the development of a short form. *Research in Developmental Disabilities, 22*(4), 273-288

Higgins, L., and Mansell, J. (2009) Quality of life in group homes and older persons' homes. *British Journal of Learning Disabilities*, 37, 207-212.

Katz, S., Ford, A.B., Moskowitz, R.W., Jackson, B.A., Jaffe, M.W. (1963) Studies of illness in the aged, The Index of ADL: a standardized measure of biological and psychosocial function. *JAMA,* 185(12), 914-9.

Katz, S., Downs, T.D., Cash, H.R., Grotz, R.C. (1970) Progress in development of the index of ADL. *Gerontologist*. 10(1), 20-30.

Kind, P., Hardman, G. and Macran, S. (1999) *UK Population norms for EQ-5D*. Centre for Health Economics, Discussion Paper, University of York, York .

Lowe, K., Felce, D., Perry, J., Baxter, H., and Jones, E. (1998) The characteristics and residential situations of people with severe intellectual disability and the most severe challenging behaviour in Wales. *Journal of Intellectual Disability Research,* 42, 375–389.

Mahoney and Barthel (1965) Functional evaluation: The Barthel Index, *Maryland State Medical Journal* 14 61–65.

Mansell, J. and Elliott, T. E. (1996) *Active Support Measure,* Tizard Centre, University of Kent, Canterbury, Kent.

Mansell, J. (2006) Deinstitutionalisation and community living: progress, problems and priorities[20]. *Journal of Intellectual and Developmental Disability*, 31,2, 65-76.

Mansell, J. and Beadle-Brown, J. (2005) Person-centred planning and person-centred action: A critical perspective. In P. Cambridge and S. Carnaby (Eds.), *Person centred planning and care management for people with learning disabilities*. Jessica Kingsley, London.

Mansell, J and Beadle-Brown, J (2005) *Engagement in Meaningful Activities and Relationship: Handbook for Observers*, Tizard Centre, University of Kent, Canterbury.

Mansell, J., Beadle-Brown, J., Whelton, R., Beckett, C. and Hutchinson, A. (2008) Effect of service structure and organisation on staff care practices in small community homes for people with intellectual disabilities *Journal of Applied Research in Intellectual Disabilities, 21*(3), 398-413.

Mansell, J and Beadle-Brown, J. (submitted) Estimating activity duration by momentary time-sampling of part or all of the day.

Ministers, Local Government, et al. (2007). Putting people first: a shared vision and commitment to the transformation of adult social care. The Stationery Office, London.

Morris, J. N., Fries, B. E., Mehr, D. R., Hawes, C., Phillips, C., Mor, V. and Lipsitz, L. A. (1994) MDS Cognitive Performance Scale*. The Journal of Gerontology,* 49: M174-M182.

Netten, A., Bebbington, A., Darton, R. and Forder, J. (2001) *Care Homes for Older People: Volume 1. Facilities, Residents and Costs*, Personal Social Services Research Unit, University of Kent at Canterbury, Canterbury.

Netten, A., Forder, J. and Shapiro, J. (2006) Measuring Personal Social Services Outputs for National Accounts: Services for Older People, PSSRU Discussion Paper 2267/3, Personal Social Services Research Unit, University of Kent, Canterbury.

Netten, A., Burge, P., Malley, J., Potoglou, D., Brazier, J., Flynn, T. and Forder, J.(2009) *Outcomes of Social Care for Adults (OSCA). Interim findings*, PSSRU Discussion Paper No.2648, Personal Social Services Research Unit, University of Kent.

Nihira, K., Leland, H., and Lambert, N. (1993) *AAMR Adaptive Behavior Scale - Residential and Community (Second Edition)*. Pro-Ed, Austin, Texas.

Qureshi, H., C. Patmore, Nicholas, E. and Bamford, C. (1998) *Overview: Outcomes of Social Care for Older People. Outcomes in Community Care Practice Series.* Social Policy Research Unit, University of York, York.

Rantz, M. and Zwygart-Stauffacher, M. (2005) A new reliable tool for researchers, regulators, providers, and consumers for measuring quality of care in nursing homes. *Nursing Outlook*, 53,4, 212.

Raynes, N. V., Sumpton, R. C., and Pettipher, C. (1989) *Index of Community Involvement,* Hester Adrian Research Centre, Manchester.

Raynes, N. V., Sumpton, R. C., and Pettipher, C. (1989) *Index of Participation in Domestic Life*, Hester Adrian Research Centre, Manchester.

Raynes, N., Wright, K., Shiell, A., and Pettipher, C. (1994). *The cost and quality of community residential care*. Fulton, London.

Rodgers, H., Curless, R. and James, O.F.W. (1993) Standardized functional assessment scales for elderly patients, *Age and Ageing*, 22, 3, 161-163.

Sen, A. (1985) *Commodities and capabilities*, Amsterdam, North-Holland.

Sen, A. (1993) Capability and well-being in M. Nussbaum and A. Sen *The Quality of Life,* Clarenden Press, Oxford.

Sloane, P., Mitchell, M., Weisman, G., Zimmerman, S., Foley, K., Lynn, M., Calkins, M., Lawton, P., Teresi, J., Grant, L., Lindeman, D. and Montgomery, R. (2002) The therapeutic environment screening survey for nursing homes (TESS-NH): An observational instrument for assessing the physical environment of institutional settings for persons with dementia, *Journal of Gerontology*, 57B, S69-S78.

Towers, A-M. (2006) *Control, Well-Being and the Meaning of Home in Care Homes and Extra Care Housing*, PSSRU Discussion Paper No. 2342, Personal Social Services Research Unit, University of Kent.

Towers, A-M., Beadle-Brown, J., Netten, A. and Smith, J. (2008) *QMF Care Homes Pilot Report*, PSSRU Discussion Paper No. 2534, Personal Social Services Research Unit, University of Kent.

University of Missouri MDS and Quality Research Team (2005) *Observable indicators of Nursing Home Care Quality Instrument* Version 10, University of Missouri, Missouri.

Welsh Centre for Learning Disabilities, Institute for Health Research, & Centre for the Economics of Mental Health (2003). *Revised Residential Services Setting Questionnaire*. Welsh Centre for Learning Disabilities, Cardiff.

Windle, K., Wagland, R., Forder, J., Janssen, D., Wistow, G. (2009) *National Evaluation of Partnerships for Older People Projects*, PSSRU Discussion Paper No. 2700, Personal Social Services Research Unit, University of Kent.

Wing, L. and J. Gould (1978) Systematic recording of behaviors and skills of retarded and psychotic children, *Journal of Autism and Childhood Schizophrenia*, 8,1, 79-97.

Wittenberg, R., Darton, R. and Netten, A. (2000) *Longitudinal Survey of People Admitted to Care Homes. Forty-Two Month Follow-up: Progress Report*, Discussion Paper No. 1621, Personal Social Services Research Unit, University of Kent at Canterbury.