

Kathy Eagar

Alan Owen

David Perkins

Linda Adamson

Karen Quinsey

Roy Harvey

Janette Green

ISBN 0 86418 659 2

April 2000

This report is part of the final evaluation report on Care Net Illawarra, one of the Australian Coordinated Care Trials based in the Illawarra. Each of the national trials addressed the same primary hypothesis:

Coordination of care of people with multiple service needs, where care is accessed through individual care plans, and funds pooled from existing Commonwealth, State and joint programs, will result in improved client health and well being within existing resources.

A summary of the key elements of the Care Net Trial is contained in Appendix 1 of this report.

There are 10 reports in this evaluation series:

Report Number 1	The Care Net Trial – What it was and How it was Managed
Report Number 2	The Care Net Intervention
Report Number 3	Care Coordination in the Care Net Trial
Report Number 4	The Use of IT in the Care Net Trial
Report Number 5	Client Experiences in the Care Net Trial
Report Number 6	The Care Net Trial – Impact on General Practitioners
Report Number 7	The Care Net Trial – Impact on Health and Community Care Providers
Report Number 8	The Care Net Trial – Impact on the Wider System
Report Number 9	The Care Net Trial – Value for Money? (this report)
Report Number 10	The Care Net Trial – The Evaluators Conclusions

Suggested citation

Eagar K et al (2000) *The Care Net Trial – Value for Money? Report 9 of the Final Evaluation of the Care Net Illawarra Coordinated Care Trial*. Centre for Health Service Development, University of Wollongong. ISBN Number 0 86418 659 2

Report Number 9

The Care Net Trial – Value for Money?

Key findings	<i>i</i>
Key findings – financial performance	<i>i</i>
Key findings – consumer outcomes.....	<i>i</i>
Key findings – conclusions.....	<i>ii</i>
9.1 Financial performance	<i>1</i>
9.1.1 Financial performance relative to the fund pool	<i>1</i>
9.1.2 Financial performance relative to the control group	<i>4</i>
9.1.3 The accuracy of the fund pool	<i>5</i>
9.1.3.1 Methodology to estimate the size of the pool	<i>5</i>
9.1.3.2 Review of the methodology	<i>6</i>
9.1.3.3 Comparison of the size of the contributions actually made with the initial and revised fund pool estimates.....	<i>6</i>
9.1.3.4 Comparison of service utilisation costs of the intervention and control groups.	<i>7</i>
9.1.3.5 Conclusion in relationship to accuracy of the pool	<i>8</i>
9.2 Outcomes for consumers	<i>9</i>
9.2.1 How health and well-being changed for intervention group clients over the period of the Trial	<i>9</i>
9.2.1.1 Intervention group clients with 2 or more assessments	<i>9</i>
9.2.1.2 Intervention group clients with 2 of more assessments at least one year apart	<i>11</i>
9.2.1.3 Comparisons over time.....	<i>12</i>
9.2.1.4 Outcomes for major diagnostic groups.....	<i>14</i>
9.2.2 Summary of intervention group outcomes	<i>16</i>
9.2.3 Outcomes Relative to the Control Group	<i>17</i>
9.2.3.1 Exits – deaths and transfers to residential aged care	<i>17</i>
9.2.3.2 Health and well-being as measured by the SF-36	<i>18</i>
9.2.3.3 Health and well-being as assessed through the Care Net assessment measures.....	<i>21</i>
9.3 How well did the Care Net assessment measures predict outcomes?	<i>24</i>
9.4 The relationship between cost and need - how well did the Care Net assessment measures predict costs?	<i>27</i>
9.4.1 The relationship between cost and exit outcome	<i>27</i>
9.4.2 The relationship between cost and need as measured by the various assessment instruments	<i>28</i>
9.4.2.1 Statistical testing of the relationship between assessment results and service costs	<i>33</i>
9.4.3 Cost of services over time	<i>34</i>
9.5 Economic costs and evaluation	<i>36</i>
9.6 Care Net – results relative to other trials	<i>37</i>
Appendix 1:An overview of the Illawarra Coordinated Care Trial	<i>38</i>
References	<i>39</i>

List of Tables

Table 1 Fund Pool Performance November 1997 to December 1999.....	1
Table 2 Administration Budget Performance November 1997 to December 1999	2
Table 3 The bottom line.....	2
Table 4 Results by funding source	2
Table 5 Fund pool deficit summary.....	3
Table 6 Differences between estimates of pool, actual contributions to pool and expenditure.	6
Table 7 Fund pool results if pool estimates had actually been applied.....	7
Table 8 Adjustments to the Care Net fund pool justified by control group costs.....	7
Table 9 Impact on fund pool justified by service use by control group	8
Table 10 Changes from 1 st to last assessment – intervention group	9
Table 11 Changes from 1 st to last assessment – long-term cohort	11
Table 12 Exits from the trial.....	17
Table 13 First Application of SF-36	20
Table 14 Second Application of SF-36	20
Table 15 Third Application of SF-36	20
Table 16 Assessment results for comparing long-term outcomes in the control and intervention groups	22
Table 17 Number of intervention clients in major exit cohorts.....	24
Table 18 Costs excluded from the Care Net bottom-line.....	36

List of Figures

Figure 1 Cost comparison - intervention and control groups.....	4
Figure 2 Changes in intervention group from 1 st to last assessment.....	10
Figure 3 Changes in long-term intervention group from trial start to trial end.....	12
Figure 4 Comparison of motor and instrumental functioning over time	13
Figure 5 Comparison of social and cognitive functioning over time.....	13
Figure 6 Comparison of self reported health status over time.....	14
Figure 7 Assessment results for different major diagnostic groups	15
Figure 8 Reasons for trial exits	18
Figure 9 Comparison of motor and cognitive function outcomes - intervention versus control groups	22
Figure 10 The assessment tools as predictors of outcome	25
Figure 11 Average cost per day by exit cohort.....	27
Figure 12 Average cost per day by blended need index score	28
Figure 13 Average cost per day by level of cognitive impairment	29
Figure 14 Average cost per day by level of domestic functioning.....	30
Figure 15 Average cost per day by level of social functioning.....	30
Figure 16 Average cost per day by self-rated health status	31
Figure 17 Average cost per day by level of motor functioning.....	32
Figure 18 Average cost per day by carer availability.....	32
Figure 19 Average cost per person per month - hospital and community care.....	34
Figure 20 Average cost per person per month - medical, diagnostic, pharmacy and emergency services	35

Key findings

Key findings – financial performance

- Care Net ended the trial \$1,805,903 (13.2%) in deficit. If administrative costs are excluded, the deficit was 12.7%.
- A review of actual contributions using the methodology initially used to establish the fund pool suggests that several contributions were above what they should have been. In particular, the PBS and DVA contributions to the pool were too high.
- If the correct contributions to the pool had been made, the final deficit would have been 25%. The initial estimates were more accurate than the contributions actually made to the pool.
- There is no evidence that intervention group members used less hospital or medical services than those in the control group.
- The control group data suggest a fund pool deficit of 14.8%. This is less than the figure of 25% quoted above because (among other factors) the method used to calculate PBS costs payments was also incorrect for the control group.
- Irrespective of the method used, the trial ended with a significant deficit. The deficit is equivalent to \$550 - \$919 per participant per year.
- The Care Net model of care coordination is not viable within current resources. The final result suggests that the financial assumptions underpinning the trial were not correct.

Key findings – consumer outcomes

- Only about 13% of the participants were 'high need' clients yet this is the group most likely to benefit from care coordination (see Report 5).
- More than 80% of participants either declined in the physical and social function (approximately 35%) or did not change (approximately 45%) over the course of the trial. Less than 20% improved.
- However, only around 20% reported that their health status had declined. Just over a half of the long-term participants in the trial actually reported that their health status had improved.
- There was no difference in death rates between the intervention and control groups.
- But more than twice as many in the intervention group were admitted to residential care than in the control group (7.5% versus 3.3%).
- There was no statistically significant difference found between the intervention and the control groups in any of the three applications of the SF-36 quality of life measure.
- There were no statistically significant differences found between intervention and control groups when the difference in scores from one application of the SF-36 to the next was tested.
- There were statistically significant differences in SF-36 scores between clients who died, those who went to residential care and those who completed the trial.
- There were no overall statistically significant differences found between the intervention and the control groups when data on motor function, cognitive function and domestic function were combined.
- Of the specific measures, only the domestic functioning results were statistically significant with the intervention group doing worse than the control group. But because of the small difference in absolute values, we do regard this as a meaningful result.
- People who exited the trial by moving to residential aged care had, on average, the worst clinical ratings at their first assessment and the worst ratings at the time of the last assessment. Their ratings were worse than the people who died during the course of the trial. Intervention group clients who completed the trial had the best ratings at both their first and their last assessment.

Key findings – conclusions

- A full economic evaluation of Care Net is not warranted. Care Net met the perceived needs of its clients and, not surprisingly, the clients were overwhelmingly positive in response. But, after two years, there is no evidence that outcomes for consumers actually improved as a result of the Care Net intervention or that the Care Net model is sustainable within current resources, however they are measured.
- The central hypothesis tested (that coordination of care of people with multiple service needs, where care is accessed through individual care plans, and funds pooled from existing Commonwealth, State and joint programs, will result in improved client health and well being within existing resources) was not supported by the evidence.

The Care Net Trial – Value for Money?

Reports 1 to 8 in this series described the Care Net Coordinated Care Trial, evaluated its processes and qualitatively evaluated the trial in terms of its impact on consumers, providers and the broader health and community care system.

This report provides the quantitative results in terms of both financial performance and consumer outcomes. Where relevant, it also makes mention of results reported in more detail in other volumes. It provides the quantitative evidence necessary to test the central hypothesis, namely that:

“...coordination of care of people with multiple service needs, where care is accessed through individual care plans, and funds pooled from existing Commonwealth, State and joint programs, will result in improved individual client health and well-being within existing resources.”

9.1 Financial performance

There are two ways to measure the financial performance of Care Net. One is relative to the fund pool. The other is relative to the control group. Each is discussed in turn. We then review the accuracy of the fund pool. In the final section of this report we discuss the need to evaluate the full economic costs and benefits of Care Net.

9.1.1 Financial performance relative to the fund pool

Table 1 summarises the performance of the Trial over the two years with respect to the fund pool. Care Net ended the trial \$1,506,169 (12.72%) in deficit¹. This table excludes the infrastructure costs of running Care Net itself which are shown in Table 2.

Table 1 Fund Pool Performance November 1997 to December 1999

Income source	income	expenditure	Difference
MBS	\$3,066,150	\$2,719,153	\$346,997
PBS	\$2,704,780	\$1,767,733	\$937,047
IAHS	\$3,679,152	\$3,527,455	\$151,697
DVA	\$1,620,422	\$1,143,386	\$477,036
Home Care	\$770,338	\$940,060	-\$169,722
HACC	\$0	\$284,667	-\$284,667
Private for profit	\$0	\$813,224	-\$813,224
Private non profit	\$0	\$167,023	-\$167,023
Care Coordination	\$0	\$1,640,511	-\$1,640,511
Contribution to administration costs	\$0	\$343,799	-\$343,799
Total	\$11,840,842	\$13,347,011	-\$1,506,169

It will be seen in the table above that the fund pool contributed \$343,799 to the overall administrative costs of running Care Net on the basis that the fund pool should carry its share of

¹ The financial data reported in this section are based on actual results as at April 2000 and are based on the trial balance at that point. The final figures are likely to vary slightly from those reported here as late invoices for the last quarter were still being received and processed. However, any variations are expected to be immaterial to the bottom line result. A further caveat is that the accuracy of the income contributed to the pool is not considered in this section (see page 5).

administrative and overhead costs. This was added to the Commonwealth grant of \$1,264,178 that was provided to run the trial and is shown in Table 2 along with other revenue sources. In addition to over-spending on the fund pool, Care Net also ended the trial with a \$299,733 deficit in its administration budget.

Table 2 Administration Budget Performance November 1997 to December 1999

Revenue Earned	Amount
Contribution Earned-Infrastructure	\$1,264,178
Interest Earned	\$161,384
Contribution from Fund Pool	\$343,799
Other Income	\$62,670
Donations	\$7,345
Total Revenue	\$1,839,376
Expenditure Incurred	Amount
Employment Costs	\$121,376
Accommodation Costs	\$90,927
Other Costs	\$835,106
Total Administration Expenditure	\$2,139,110
Net total	-\$299,733

The bottom line for the trial is summarised in Table 3. Care Net ended the trial with a total cash deficit of \$1,805,903. This represents 13.2% of total income (both fund pool and administration).

Table 3 The bottom line

Fund Type	Income	Expenditure	Deficit	% deficit
Fund pool (table 1)	\$11,840,842	\$13,347,011	-\$1,506,169	-12.7%
Admin fund (table 2)	\$1,839,376	\$2,139,110	-\$299,734	-16.3%
Total	\$13,680,218	\$15,486,121	-\$1,805,903	-13.2%

The bottom line result is significantly worse than the projection we included in our mid-term report. At that time, we projected that the deficit would be between 5.6% to 10.5% of total income.

Table 4 shows results for the fund pool by funding source. It will be seen that the trial made considerable savings on Commonwealth funded services, utilising only 76% of Commonwealth funding for the services for which the funds were pooled. The expenditure of State funds was much closer to actual income (7% variance). These results raise the obvious question of whether the contributions to the pool were correctly calculated. This issue is discussed on page 5.

Table 4 Results by funding source

	income	expenditure	difference	% saving
Commonwealth funded (MBS, PBS, DVA)	\$7,391,352	\$5,630,272	\$1,761,080	23.83%
State funded (IAHS, HACC and Home Care)	\$4,449,490	\$4,752,182	-\$302,692	-6.80%
Other (Private, care coordination and other)	\$0	\$2,964,557	-\$2,964,557	0.00%
All	\$11,070,504	\$12,406,951	-\$1,336,447	-12.72%

Table 5 summarises the bottom line with respect to the hypotheses underpinning the trial. Assuming that the pool contributions were correct (see page 5), Care Net saved \$478,141 in the costs of services that were pooled. However, Care Net spent nearly \$4 in new community services, care coordination and associated costs for every \$1 that it saved in hospital and medical costs.

Table 5 Fund pool deficit summary

	Amount
Savings in service costs	\$478,141
Spending on care coordination and other costs	-\$1,984,310
Total	-\$1,506,169

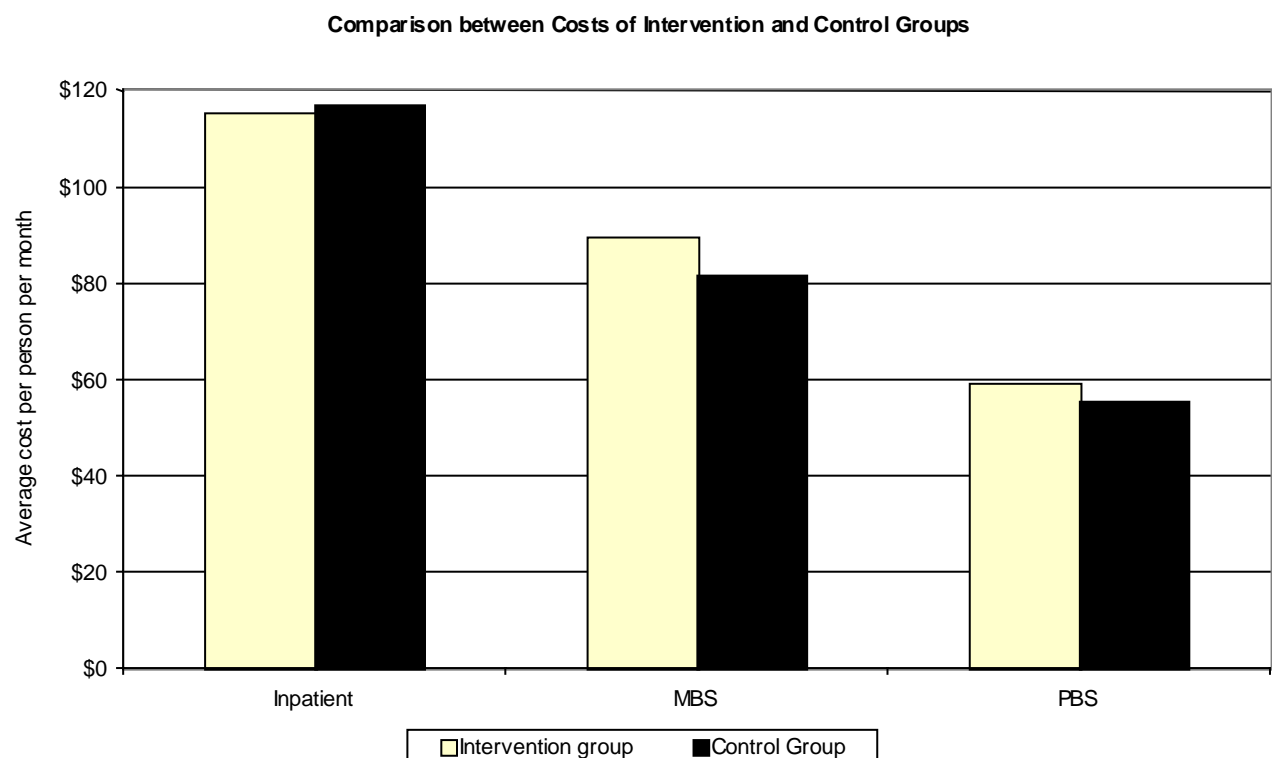
This raises an important issue. Close to 17% of the Care Net fund pool was spent on Care Coordinators and associated costs. The trial could only have been viable if services costing nearly \$2 million had been reduced and if the \$2 million 'saved' had been re-directed to offset the costs of care coordination. While there is anecdotal evidence suggesting some level of 'waste' in current service provision (unnecessary interventions, inefficiencies and so on), there is little reason to suggest that close to 17% of existing resources are currently being wasted.

Even if this were the case, there is little reason to suggest that a trial such as Care Net had any possibility of eliminating that waste. Our conclusion at the midterm was that the Care Net model of care coordination was not viable within current resources. The final result reinforces that conclusion and suggests that the financial assumptions underpinning the trial were simply not correct.

9.1.2 Financial performance relative to the control group

Figure 1 summarises the average monthly cost of services by individuals in the intervention and control groups. The figures shown are on a per capita basis and have been adjusted to take account of differences in the size of the intervention and control groups as the trial progressed. Only three service costs are included as these are the only three for which reliable control group data are available. Nevertheless, these three represent the majority of incurred costs.

Figure 1



As Figure 1 graphically illustrates, the cost per person is remarkably similar between the two groups – a difference of only \$2 per person for inpatient care, \$7 for MBS funded services and \$4 for PBS funded pharmaceuticals. Across the 26 months of the trial, average costs for these 3 services for the intervention group were \$248 more per person than the average costs for the control group.

We are cautious about using the control group as a comparator for PBS costs. As discussed in Reports 4 and 6, General Practitioners participating in Care Net used a software program known as 'Medical Director' to manage their prescribing of pharmaceuticals. Our GP survey found that 77% of GPs reported that it assisted with prescribing for intervention group patients. But 75% also reported that it assisted with prescribing for control group patients, suggesting that the control group results have been 'contaminated' by the control group having access to an intervention designed specifically for the intervention group.

In summary, after two years of operation, there was no evidence that intervention group members used less hospital or medical services than those in the control group. Reliable comparisons on pharmaceuticals are not available due to the contamination of the data.

9.1.3 The accuracy of the fund pool

There are three methods that can be used to assess the accuracy of the funds that were contributed to form the Care Net fund pool:

- Review the methodology used to estimate the size of the pool
- Compare the size of the contributions actually made with the above estimate and
- Compare service utilisation costs of the intervention and control groups.

Each of these approaches is discussed in turn.

9.1.3.1 Methodology to estimate the size of the pool

Prior to taking on the role as local evaluators, the CHSD was engaged by Care Net to estimate the size of the fund pool. The method used to construct the pool has been described elsewhere (Cromwell et al., 1997, Cromwell et al 1999).

The estimate of the HIC contribution for PBS services was revised in September 1998 (Cromwell, 1998). This occurred because the initial estimation of the HIC pool contribution for PBS services was based on tracking data that were supplied using a set of data extraction rules that subsequently changed in the live phase of the trial. The change in the extraction rules was associated with the PBS safety-net cards. These cards can apply to a single person or a household, and are allocated when the person or household have spent more than a certain amount on prescription drugs. When more than one person is covered by the card, the HIC cannot identify which scripts belong to which person. Consequently, because the original tracking data included all PBS utilisation information associated with a client's safety-net card, the supplied data included PBS information on people not in the trial but who were covered by the safety-net card of a client.

The HIC changed the rules governing the extraction of data for the live phase of the trial. Only PBS utilisation data that could be attributed directly to a client in the trial were released. So, once a client started to use a safety-net card, data were only extracted for clients who were the sole person registered on the safety-net card or who the HIC had been able to identify because the client supplied their Medicare card at the time they collected the prescription.

This change effected the initial pool estimate and HIC contribution. The original estimate was based on the utilisation data that included information on people not in the trial. However, the changes meant that the rules under which funds were contributed and paid back were no longer equivalent. With the HIC only billing for identifiable clients, they would only receive back a portion of the contributed funds. Therefore, tracking data extracted using the new rules were requested from the HIC and a pool estimate consistent with those rules was derived.

The revised forecasts based on the new data and rules were around 72% of the old estimate.

A second revision of the IAHS contribution to the pool was then prepared in February 1999 based on the actual profile of people participating in the trial (Cromwell, 1999). The initial estimate had been based on incomplete data from the tracking phase.

9.1.3.2 *Review of the methodology*

The methodology used by the CHSD to estimate contributions from the Illawarra Area Health Service was independently reviewed by Acumen Alliance at the request of the Commonwealth. The Acumen Alliance review agreed with the methodology used to calculate the pool and concluded that the CHSD method 'resulted in the most appropriate capitation rates'. However, at the time of writing, the data used by Acumen Alliance had not been reconciled with the data used by the trial and the local evaluators.

While the trial undertook some internal analysis of the other funding sources, they were not subjected to independent review.

9.1.3.3 *Comparison of the size of the contributions actually made with the initial and revised fund pool estimates*

While the size of the pool was estimated by the CHSD, the CHSD figures were not necessarily used to determine the actual contribution made to the pool. Rather, the Project Manager undertook various negotiations with funding authorities as the trial proceeded and some of these resulted in changes to the pool.

Significantly, while a new pool estimate was calculated for PBS services, no revision was made to the contribution actually made for PBS services. Conversely, the DVA contribution was increased without any external review.

Table 6 compares the difference between what would have been contributed to the pool if the CHSD estimates had been applied to all funding sources (Column 2) with what was actually provided (Column 3). The percentage difference between these two is shown in Column 9. It will be seen that the IAHS contributed exactly what had been estimated and that the MBS contribution was close to the estimate (4.2% difference). However, both the PBS and the DVA contributions were significantly larger (32% and 29% respectively) than had been estimated.

Table 6 Differences between estimates of pool, actual contributions to pool and expenditure.

Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8	Column 9
Fund source	CHSD estimate of pool size	Actual pool	Actual expenditure	Difference between CHSD estimate and expenditure	Difference between actual pool and expenditure	Difference between CHSD estimate and actual pool	% difference b/n CHSD estimate and actual expenditure	% difference between CHSD estimate and actual pool
MBS	\$2,941,614	\$3,066,150	\$2,719,153	\$222,461	\$346,997	\$124,536	7.6%	4.2%
PBS	\$2,043,857	\$2,704,780	\$1,767,733	\$276,124	\$937,047	\$660,923	13.5%	32.3%
IAHS	\$3,679,152	\$3,679,152	\$3,527,455	\$151,697	\$151,697	\$0	4.1%	0.0%
DVA	\$1,259,990	\$1,620,423	\$1,143,386	\$116,604	\$477,037	\$360,433	9.3%	28.6%
Total	\$9,924,612	\$11,070,505	\$9,157,727	\$766,885	\$1,912,778	\$1,145,893	7.7%	11.5%

Column 5 shows the difference between the CHSD estimate and the actual trial expenditure. These differences are shown as percentages in Column 8. In all cases, the CHSD estimate was above the actual level of expenditure. The closest estimate was for IAHS services (4.1% difference between estimate and actual expenditure). While the estimate for inpatient services was under, the estimate for outpatient and community services was over. The biggest difference was between the CHSD estimate and PBS expenditure (14%) followed by the DVA estimate (9% difference).

Importantly, the two income sources where the trial made the most profit (see Table 1) are the two sources who contributed considerably more to the pool than the CHSD estimate had suggested was required. These two sources are the PBS and the DVA, both of which are Commonwealth funded.

Table 7 shows the effect on the pool if the initial methodology for calculating the pool had been applied. While the trial would still have saved against MBS, PBS, IAHS and DVA, the net result is a deficit of \$2.65 million or 25%.

Table 7 Fund pool results if pool estimates had actually been applied

Income source	income	expenditure	Difference
MBS	\$2,941,614	\$2,719,153	\$222,461
PBS	\$2,043,857	\$1,767,733	\$276,124
IAHS	\$3,679,152	\$3,527,455	\$151,697
DVA	\$1,259,990	\$1,143,386	\$116,604
Home Care	\$770,338	\$940,060	-\$169,722
HACC	\$0	\$284,667	-\$284,667
Private for profit	\$0	\$813,224	-\$813,224
Private non profit	\$0	\$167,023	-\$167,023
Care Coordination	\$0	\$1,640,511	-\$1,640,511
Contribution to administration costs	\$0	\$343,799	-\$343,799
Total	\$10,694,951	\$13,347,011	-\$2,652,060

9.1.3.4 Comparison of service utilisation costs of the intervention and control groups.

The previous analysis that compared the costs of the intervention and control group provides the third method to assess the accuracy of the Care Net fund pool. The results are shown in Table 8. If the Care Net pool had been based on actual spending by the control group over the same period, the size of the pool would have \$273,422 less than it actually was. This would increase the size of the fund pool deficit to \$1,712,154 or 14.8%.

Table 8 Adjustments to the Care Net fund pool justified by control group costs.

	Inpatient	MBS	PBS	Average number in trial	Total
Intervention client per month	\$115	\$89	\$59	1,104	
Control client per month	\$117	\$82	\$55	576	
Difference	\$2	-\$8	-\$4		
Change justified per intervention client over trial	\$49.82	-\$197.52	-\$99.86		-\$247.56
Change justified to total Care Net pool	\$55,025	-\$218,149	-\$110,297		-\$273,422

Based on spending on the control group, the IAHS contribution would be increased by \$55,025 while the MBS and PBS contributions would have been decreased by \$218,149 and \$110,297 respectively.

The effects of making such changes are shown in Table 9. For the MBS, the outcome would be a 5% saving against the pool. For the PBS, the outcome would be a significant saving of 32% against the pool. For inpatient services provided by the IAHS, the effect would have been that Care Net utilised \$152,940 (5%) more IAHS services than IAHS contributed to the pool.

Table 9 Impact on fund pool justified by service use by control group

	Adjusted income	Actual Expenditure	Difference	% saving
MBS	\$2,848,001	\$2,711,238	\$136,764	4.80%
PBS	\$2,594,483	\$1,764,057	\$830,426	32.01%
Inpatient	\$3,046,477	\$3,201,257	-\$154,780	-5.08%

As we have already suggested, we believe that there was some 'contamination' between the control and client participants. Doctors could have patients assigned to both the control and the intervention groups. The doctor, however, could use Medical Director for both groups and would have access to e-mail facilities for both. The 'culture' of the trial could have influenced the treatments offered by general practitioners to both groups of participants.

9.1.3.5 Conclusion in relationship to accuracy of the pool

The results suggest that there is reason to adjust the size of the pool. However, we are not suggesting that such adjustments be made at this point. Rather, the bottom-line for Care Net does not substantially change except to the extent of the deficit:

- The trial fund pool ended with an actual cash deficit of 12.7%. If administrative costs are included, the total deficit for the trial was 13.2%.
- The review of actual contributions relative to the pool estimate suggests a deficit of 25%.
- The control group data suggest a deficit of 14.8%.

Irrespective of the method used, the trial ended with a significant deficit. That deficit is equivalent to \$1,260 - \$1,991 per intervention client over the course of the trial or \$550 - \$919 per participant per year.

9.2 Outcomes for consumers

Financial results are but one measure of the outcomes of the Care Net trial. While Care Net could not be delivered within existing resources, it is possible that the financial costs were justified by improved consumer outcomes. In this section we report on whether this was the case. Two measures were used to assess the overall impact of the trial as an intervention on participants' health and well-being:

- Assess the trial clients at the beginning, mid-point and end of the trial using SF-36 data and data from the trial assessments.
- Compare the trial clients (the intervention group) with the trial controls.

Each of these approaches is discussed in turn.

9.2.1 How health and well-being changed for intervention group clients over the period of the trial

In our mid term report we described each of the measures used to assess trial participants and provided a range of descriptive and analytic statistics to show any impact of the trial at that point. Rather than repeat that description and analysis here, readers are referred to our first report (Eagar et al., 1998).

Instead, this section simply summarises key results.

9.2.1.1 Intervention group clients with 2 or more assessments

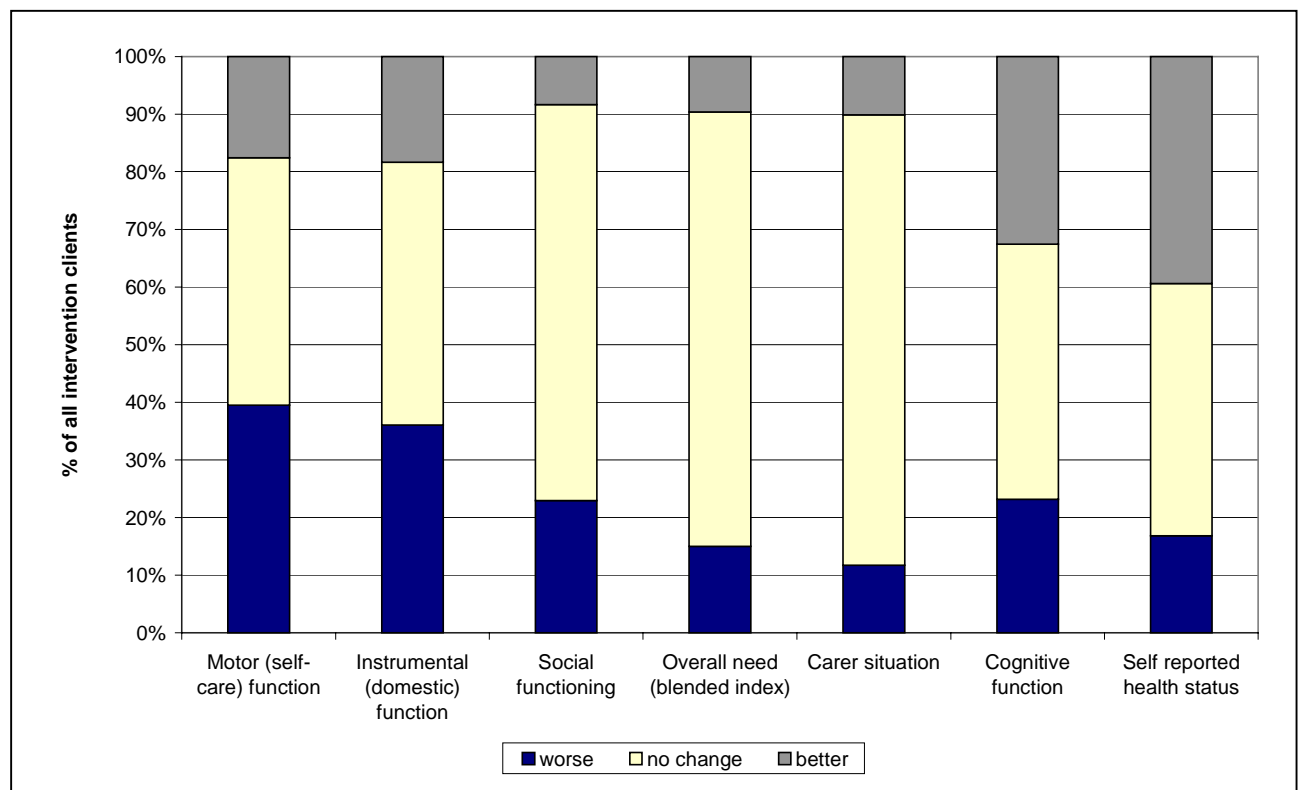
A total of 1232 intervention group clients had 2 or more assessments. Their results on 7 key measures are show in Table 10.

Table 10 Changes from 1st to last assessment – intervention group

Measure	Number				Percentage			
	worse	no change	better	total	worse	no change	better	total
Motor (self-care) function	466	506	207	1179	39.5%	42.9%	17.6%	100.0%
Instrumental (domestic) function	425	537	216	1178	36.1%	45.6%	18.3%	100.0%
Social functioning	269	805	98	1172	23.0%	68.7%	8.4%	100.0%
Overall need (blended index)	185	928	119	1232	15.0%	75.3%	9.7%	100.0%
Carer situation	135	899	117	1141	11.8%	78.8%	10.3%	100.0%
Cognitive function	273	522	384	1179	23.2%	44.3%	32.6%	100.0%
Self reported health status	167	434	391	992	16.8%	43.8%	39.4%	100.0%

These results are shown graphically in Figure 2.

Figure 2 Changes in intervention group from 1st to last assessment



About half of the intervention group did not change between their first and last assessments. However, there were some important differences between the dimensions.

The availability of a carer remained the most stable attribute measured through the trial, with nearly 80% of participants having the same situation at their last assessment as they did at the first one.

This was followed by the measure of overall need (the blended needs index), with about 75% of participants recording the same level at the last assessment as they did at the first one.

Of the clinical and psychosocial measures, social functioning remained the most stable. Nearly 70% of participants recording the same score at the last assessment as they did at the first one.

Over 50% of participants changed in their self-reported health status or in their motor, instrumental or cognitive functioning during the two years of the trial. Instrumental (domestic) and motor function were the dimensions showing the most decline, with nearly 40% of all participants declining on these measures since their first assessment. Less than 20% of the cohort improved on these two measures.

The pattern for self-reported health status and cognitive functioning differed from that of the physical measures. Nearly 40% self-reported an improvement in their health status while only 17% reported a decline.

Likewise, over 30% of all participants showed an improvement in cognitive functioning and 16% a decline. However, there is concern about the validity of the cognitive results. The same cognitive test was repeated throughout the trial (some participants were tested more than 6 times) and the same questions were asked on each occasion. For example, one question required respondents to recall a name and address at the end of the test. It appears that the results are at least partially

attributable to a learning effect in that some respondents had plenty of opportunity to ‘learn the right answer’. As one respondent commented:

“I used to practise the address (a component of the cognitive scale) all morning before the care coordinator came.”

9.2.1.2 Intervention group clients with 2 of more assessments at least one year apart

The problem with the figures shown above are that they include participants who were only in the trial for a short period and participants whose first and last assessments were very close together (in some cases, only a matter of days).

The control group were only assessed at three points – around November 1997 (trial start), around November 1998 (mid point) and after the end of the trial whereas the intervention group were assessed more frequently. A valid comparison of the intervention and the control groups requires that the control group and the intervention group be compared at similar points in time.

Accordingly, the next comparison uses only intervention data in which the participant had:

- Their first assessment before September 1998 and
- Their last assessment after June 1999.

440 intervention group members (34%) met these criteria and they are termed the ‘long term cohort’ in Table 11 and Figure 3.

Table 11 Changes from 1st to last assessment – Long-term cohort

Measure	Number				Percentage			
	worse	no change	better	total	worse	no change	better	total
Motor (self-care) function	196	162	82	440	44.5%	36.8%	18.6%	100.0%
Instrumental (domestic) function	189	160	90	439	43.1%	36.4%	20.5%	100.0%
Social functioning	125	266	45	436	28.7%	61.0%	10.3%	100.0%
Overall need (blended index)	66	354	20	440	15.0%	80.5%	4.5%	100.0%
Carer situation	69	293	58	420	16.4%	69.8%	13.8%	100.0%
Cognitive function	100	182	158	440	22.7%	41.4%	35.9%	100.0%
Self reported health status	66	116	188	370	17.8%	31.4%	50.8%	100.0%

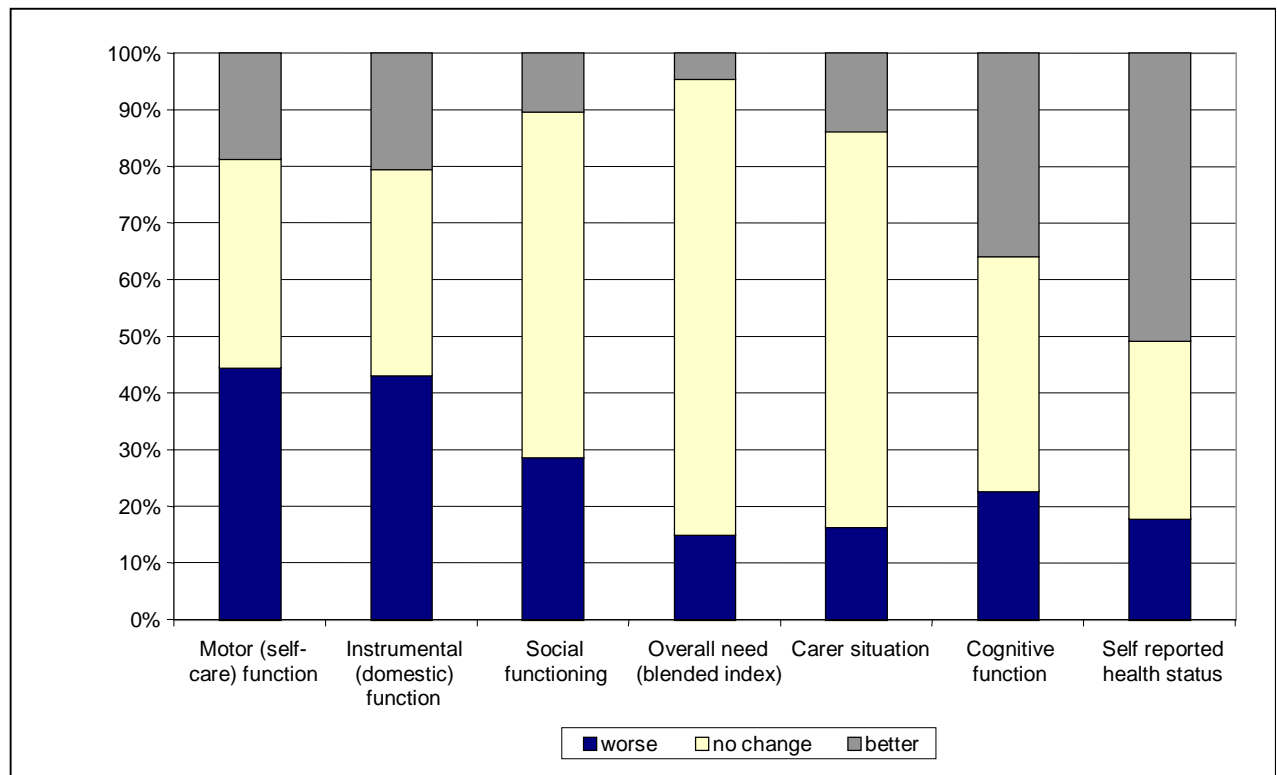
This long-term cohort has a different profile than the total group described above. The availability of a carer was less stable than the group as a whole, with 70% (compared to 80% of all intervention participants) having the same situation at their last assessment as they did at the first one.

However, overall need (the blended needs index) was more stable, with over 80% of this cohort recording the same level at the last assessment as they did at the first one.

Of the clinical and psychosocial measures, social functioning was again the most stable. 61% (compared to 69% for the total group) recording the same score at the last assessment as they did at the first one.

Not surprisingly, over 60% of the long-term cohort changed in their self-reported health status or in their motor, instrumental and cognitive functioning during the two years of the trial. Instrumental (domestic) and motor function showed more decline but, importantly, self-reported health status and cognitive functioning showed more improvement.

Figure 3 Changes in long-term intervention group from trial start to trial end



9.2.1.3 Comparisons over time

Figure 4 shows a comparison of the results of the two key physical function measures at different points in the trial. The mid-term results are those we reported in December 1998. It will be seen that:

- Relative to the mid-term results, there is more change at the end of the trial. This is not surprising, given that participants have had more opportunity (time) to change.
- The long-term cohort shows the most change, with most recorded change being a change for the worse rather than the better.

These findings are consistent with what had been expected of a cohort of the frail, elderly population.

Figure 4 Comparison of motor and instrumental functioning over time

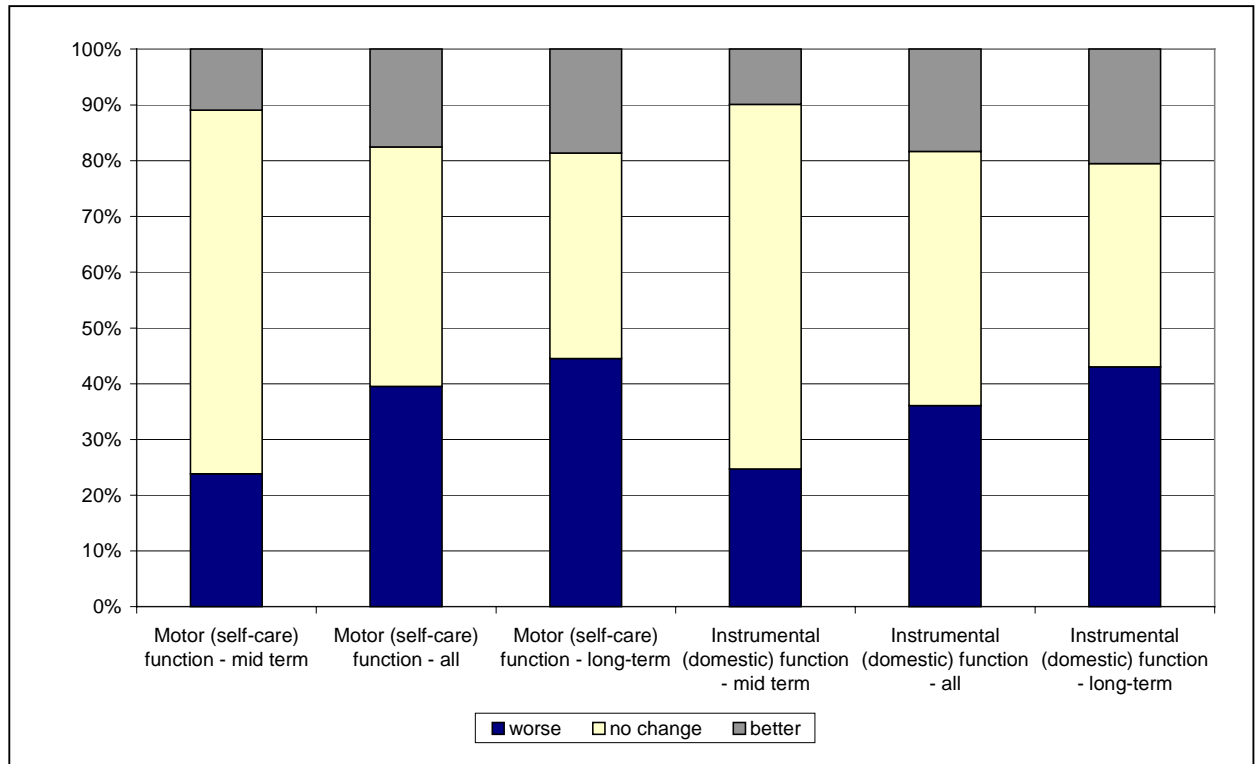
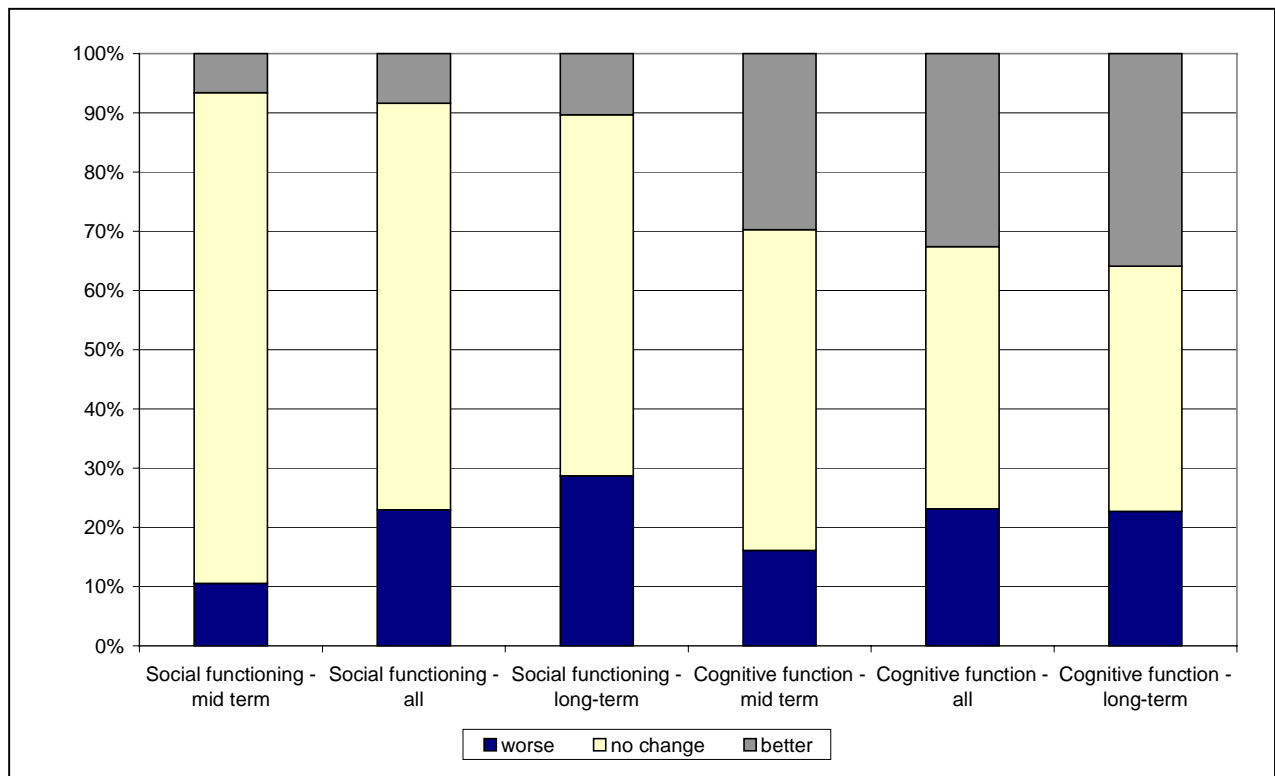


Figure 5 Comparison of social and cognitive functioning over time

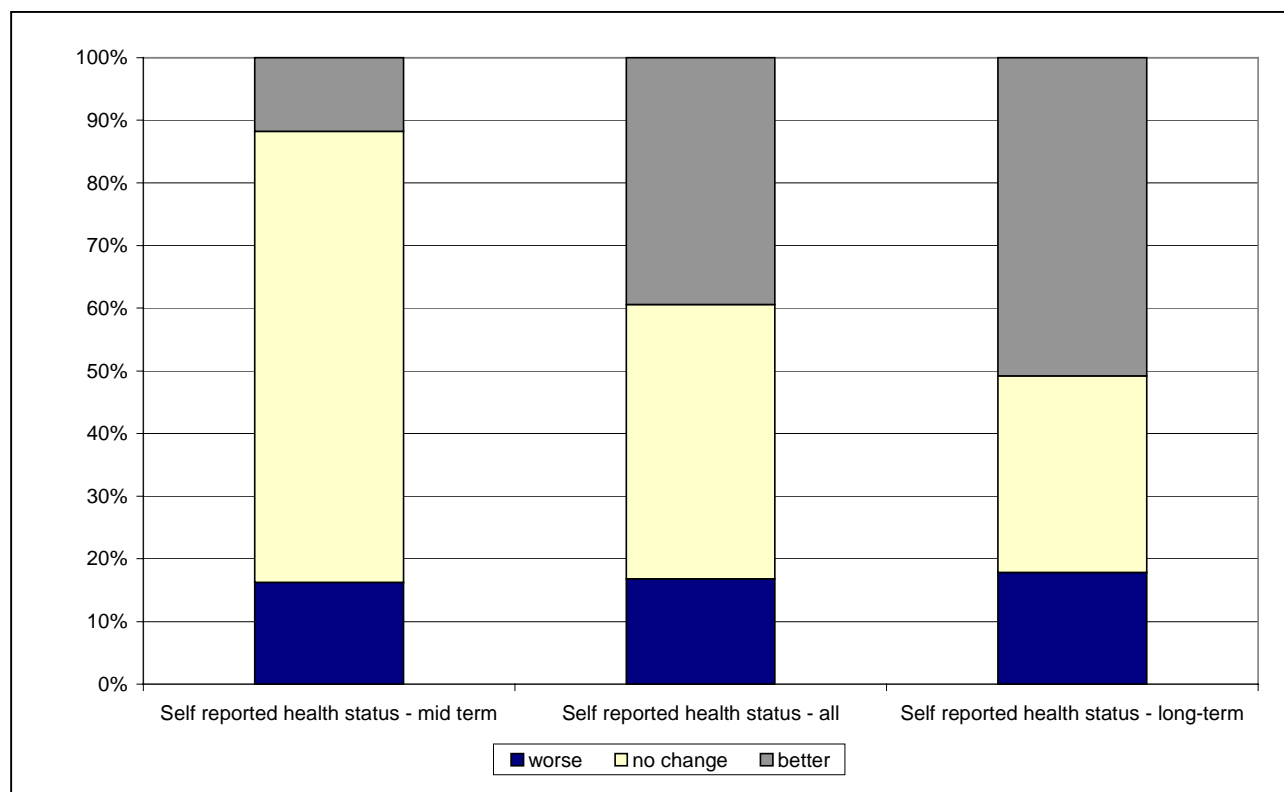


The results for the physical measures are different from the results for cognitive functioning shown in Figure 5. As before, there is more change at the end of the trial than there was at the mid-point. However, it will be seen that, for cognitive function, the long-term cohort showed the most improvement, with about 36% improving and only 23% declining.

Figure 5 also shows results on the social functioning scale. While more stable than the other measures, it’s pattern follows the pattern for physical functioning.

Importantly, self-reported health status shows significant differences over time as illustrated in Figure 6. At the end of the first year of the trial, 72% of participants reported no change, with 12% reporting a change for the better and 16% reporting a change for the worse. This contrasts with the result for the long-term cohort with only 31% reporting no change. Only 18% reported a decline and, importantly, just over half (51%) reported an improvement in their health status over the two years of the trial.

Figure 6 Comparison of self reported health status over time



9.2.1.4 Outcomes for major diagnostic groups

Figure 7 shows the results of the first and last assessments for 4 of the major diagnostic groups included in the trial – circulatory diseases, diabetes, musculoskeletal conditions and respiratory conditions².

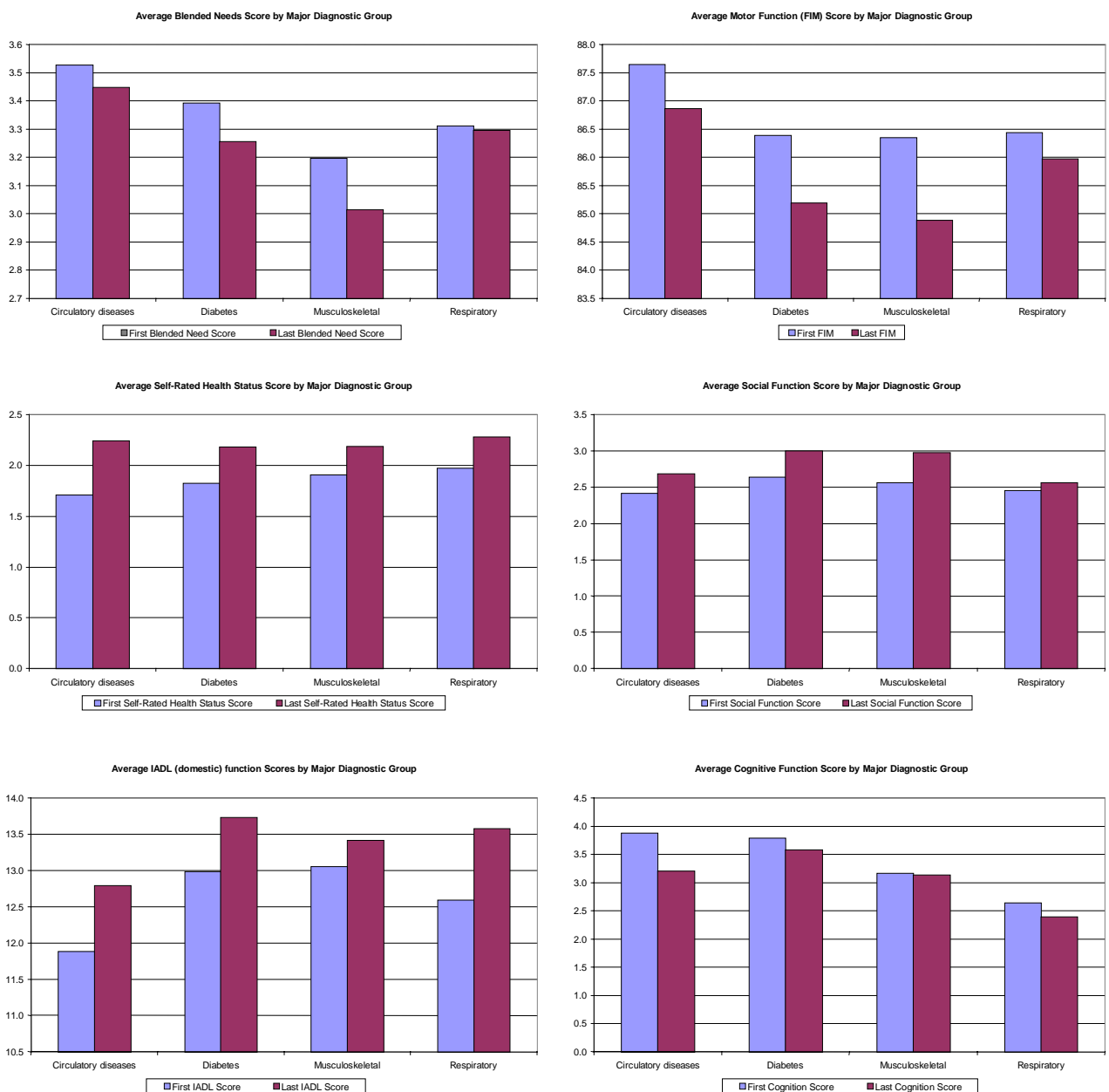
Intervention group clients were classified to one of these major diagnostic groups based on the first diagnosis recorded in their medical record. As already noted, an average of 4 diagnoses was recorded for each person but the following analysis is based only on the first recorded diagnosis. As such, it is equivalent to an analysis of a ‘principle diagnosis’ in the acute hospital setting. Clients with conditions not included in these 4 diagnostic groups were excluded from the analysis.

² See Report 5 for a description of how diagnostic data were collected and for a discussion as to its accuracy.

In general, clients with diabetes or musculoskeletal conditions had worse ratings than those clients with circulatory diseases or respiratory conditions. However, there were some exceptions. Clients with respiratory conditions rated their health as worse than any of the other groups while clients with circulatory diseases and diabetes had lower cognitive function than the other groups.

The outcomes for each group are about the same on most measures, although the musculoskeletal group declined a little more than the others on five of the six measures. The only exception is for the IADL measure, where the circulatory and respiratory groups experienced the most decline.

Figure 7 Assessment results for different major diagnostic groups



9.2.2 Summary of intervention group outcomes

In summary, the overall conclusion is that:

- Only about 13% of the participants were ‘high need’ clients yet this is the group most likely to benefit from care coordination (see Report 5).
- More than 80% of participants either declined in the physical and social function (approximately 35%) or did not change (approximately 45%) over the course of the trial. Less than 20% improved.
- However, only around 20% reported that their health status had declined. Just over a half of the long-term participants in the trial actually reported that their health status had improved.
- Around one third of participants showed an improvement in the cognitive functioning, with around one quarter showing a decline, but some of this effect is probably due to the repeated use of one measure throughout the trial.

9.2.3 Outcomes Relative to the Control Group

As noted in Report 3 the care coordinators collected standardised measures on the intervention group. These same measures were collected at three points for the control group – at the start of the trial, at the mid-point and at the trial conclusion³.

As we have previously noted, there was some ‘contamination’ between the control and client participants. It should also be noted that the last collection of the SF-36 data may reflect some participants’ feelings about the closure of the trial – as well as, or instead of, their more general outlook on life.

9.2.3.1 Exits – deaths and transfers to residential aged care

One of the most basic outcome measures relevant to the Care Net trial is ‘disposition’ – the number of people who died or who went to residential care during the 26 months of the trial. Table 12 shows reasons for exits from the trial. In total, 549 participants (27.62%) left the trial before it was concluded. More control group participants left the trial than intervention group participants (35.23% versus 23.52%). The trial entered its wind-down phase in May 1999 and this is reflected in the declining participant figures from that date. Participants originally gave consent to participate in the trial for the period ending June 1999. Re-consent was required for the period July to December 1999 and, not surprisingly, significantly more control group members elected to not give consent to remain in the trial for the extension period (July 1999 to November 1999).

Table 12 Exits from the trial

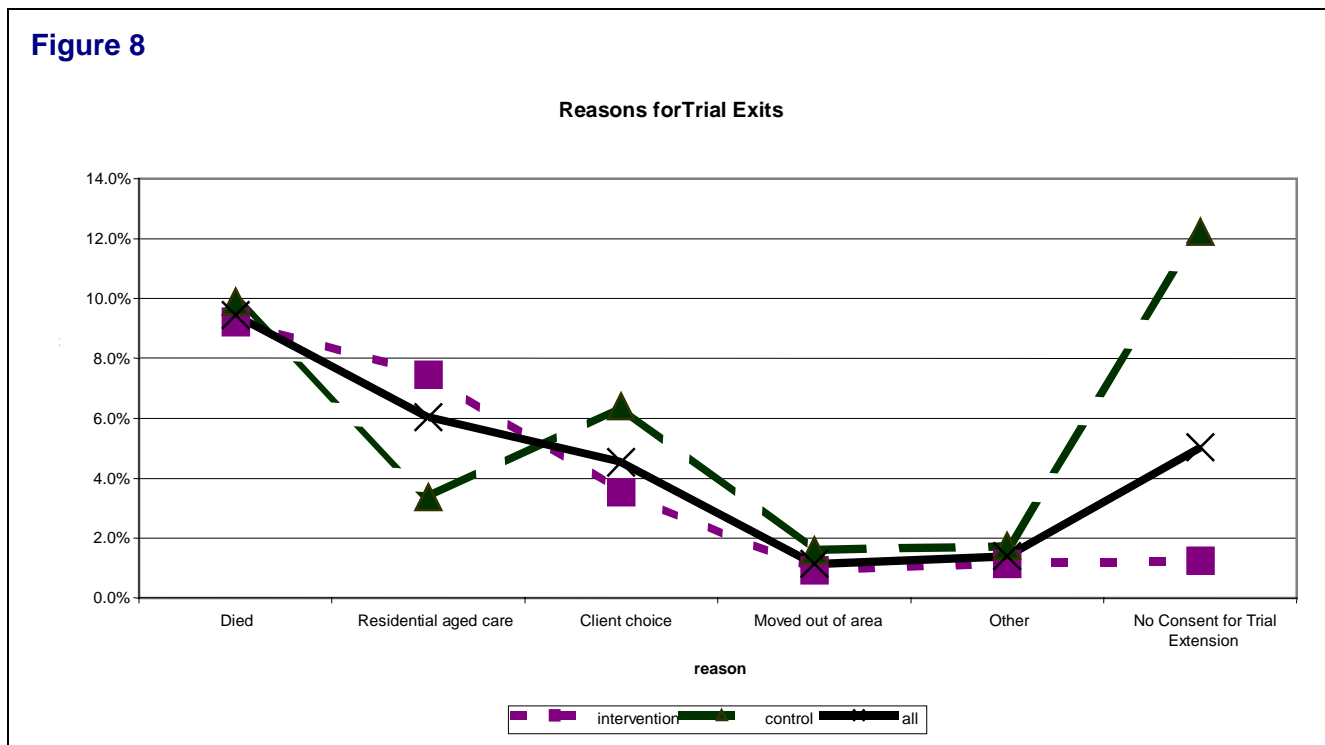
	Intervention	Control	All
Died	120	68	188
Residential aged care	97	23	120
Client choice	46	44	90
Moved out of area	12	11	23
Other	15	12	28
No Consent for Trial Extension	16	84	100
All Withdrawals	306	242	549
Total withdrawals excluding consents	290	158	449
Total in trial at trial end	995	445	1439
Total in trial at trial start	1301	687	1988
Died	9.2%	9.9%	9.5%
Residential aged care	7.5%	3.3%	6.0%
Client choice	3.5%	6.4%	4.5%
Moved out of area	0.9%	1.6%	1.2%
Other	1.2%	1.7%	1.4%
No Consent for Trial Extension	1.2%	12.2%	5.0%
Exits as % of all participants	23.52%	35.23%	27.62%
% excluding ‘no consent’	22.29%	23.00%	22.59%

A summary of the results reported above is shown graphically in Figure 8. Approximately 10% of participants died during the course of the trial, with no significant difference between the intervention and control groups. A further 6% were admitted to a residential care facility.

The key, and statistically significant, result is that more than twice as many in the intervention group were admitted to residential care than in the control group (7.5% versus 3.3%).

³ Due to a decision taken by the trial management the collection of final assessment data for the control group was cancelled in the last months of the trial. However, the Management Committee subsequently overturned this decision. One consequence was that the assessment data for the control group were not collected until February and March 2000 – after the conclusion of the trial and several months after the last assessment of intervention clients.

Figure 8



9.2.3.2 Health and well-being as measured by the SF-36

Background on the SF-36 and how it was used in the trial

The SF-36 Health Survey was designed to measure health-related well-being and quality of life. When it is possible, the client completes the survey. As its name suggests, the survey consists of 36 items that are designed to provide information as to how the patient’s health status and level of physical independence affect different aspects of his or her life.

One item stands alone. This question asks the client how they would rate their health compared to one year ago.

Responses on the remaining 35 items can be grouped into eight subscales – Physical Functioning, Role-Physical, Bodily Pain, General Health, Vitality, Social Function, Role-Emotional and Mental Health.

A further summary is possible if these eight scores are combined to form just two underlying component scores which have been identified in a number of different data sets when Factor Analysis was applied to the eight SF-36 subscales. These two component scores are the physical component score (PCS) and the mental component score (MCS). To calculate these, the subscale scores are standardised using means and standard deviations from the underlying population, which is the Australian population in this case. The standardised scores are then aggregated using the factor score coefficients obtained for the relevant population. These aggregated values are then standardised so that, if they had been calculated for the underlying population, they would have a mean of 50 and a standard deviation of 10. Because of this standardisation, a group such as the set of trial participants can be compared with the Australian population.

In the Illawarra Coordinated Care Trial there were three administrations of the SF-36 in both the control and the intervention groups. Each of the eight subscale scores was calculated for each

client. These subscale scores were then combined into PCS and MCS scores, using means, standard deviations and factor score coefficients from the Australian population in the calculations. Analysis has been undertaken on the two component scores, PCS and MCS, as well as the eight subscales, as both perspectives provide interesting information.

Analysis method

We investigated whether the intervention improved the general health and well-being of the trial clients by testing for differences between the intervention and the control groups at each application of the SF-36. We also examined the intervention group to see if there was an increase in scores from one application to the next. If we find an increase, we could then test whether it was larger than any increase in scores in the control group. If there was no increase in scores in the intervention group, we could test whether the deterioration is less in the intervention group than in the control group, or whether a significantly larger number of people improved in the intervention group than in the control group.

When there is interest in comparing groups on a number of variables, a multivariate analysis of variance (MANOVA) is an appropriate technique for statistical analysis. We used it to compare the control and intervention groups on all eight subscales together, and, alternatively, the two component scores together. Any differences can be more clearly understood by examining the differences on individual variables, when it is of interest or appropriate to do so. Statistical testing was carried out at a 5% significance level.

SF-36 results - PCS and MCS scores

Correlations between the component scores and the eight subscale scores fell into the expected pattern. For the PCS, correlations were highest with the Physical Functioning, Role-Physical and Bodily Pain subscales. The correlations between these subscales and the MCS were very small. On the other hand, the Mental Health, Role-Emotional and Social Functioning subscales correlated well with the MCS, but poorly with the PCS. General Health and Vitality correlated moderately well with both component scores.

Means and standard deviations for each of the component scores were calculated for all three sets of SF-36 data. Details are provided in the three tables following. At each application of the instrument and for each group, the average of the PCS is well below that of the general Australian population, which has a mean of 50. The MCS averages are also lower than the population average, but the difference is much less.

It is clear from the three tables below that the standard deviations are quite close to 10. However, at the first application of the SF-36, there was less variability in the data, resulting in a standard deviation of roughly 8 in both the intervention and the control groups. At the mid-term and the final applications of the SF-36, the standard deviation had increased to a little over 11 for both intervention and control groups.

Table 13 First Application of SF-36

Group	Average PCS	Stand Dev PCS	Average MCS	Stand Dev MCS	Number of Clients
Intervention	33.76	11.41	44.01	8.17	1221
Control	34.31	11.28	44.05	8.36	631

Table 14 Second Application of SF-36

Group	Average PCS	Stand Dev PCS	Average MCS	Stand Dev MCS	Number of Clients
Intervention	33.59	11.15	47.15	12.00	990
Control	32.97	11.04	47.91	11.39	482

Table 15 Third Application of SF-36

Group	Average PCS	Stand Dev PCS	Average MCS	Stand Dev MCS	Number of Clients
Intervention	32.17	10.73	45.65	11.66	871
Control	32.43	10.47	46.54	11.62	377

There was no statistically significant difference found between the intervention and the control groups in any of the three applications of the SF-36. Moreover, no overall statistically significant differences were found between intervention and control groups when the difference in scores from one application of the SF-36 to the next were tested. This means that, considering all the variables together, there was no indication that the intervention and control groups were different with respect to the change in SF-36 scores from one application to the next.

However, if the subscales were tested one by one, some were found to increase between applications and some to decrease, while the rest showed no real change at all. Mental Health showed the strongest improvement between the beginning and the midpoint of the trial for both the intervention group and the controls, but this pattern was not repeated after the midpoint: Vitality decreased more than any other subscale during the first half of the trial. After this, it showed a further decline, but several other subscales decreased by larger amounts during the second half of the trial. In fact, during this later part of the trial, all subscales decreased.

Various patterns amongst the subscales were also observed when the differences between applications were investigated.

- General Health, Physical Functioning and Social Functioning increased significantly more for the intervention clients than for the controls between the 1st and 2nd applications of the instrument,
- Physical Functioning declined more for the control group than for the intervention clients between the 1st and 3rd administrations of the SF-36, and
- Mental Health declined more for the control than for the intervention group between the 2nd and 3rd applications.

When the component scores were investigated separately, only one significant difference was found. The PCS declined more for the control than for the intervention group between the 1st and 2nd applications of the SF-36 but this statistical difference disappeared between the 2nd and 3rd applications. The MCS improved over the same time period, but not significantly more for one group than for the other.

In summary, there were no statistical differences between the intervention and control groups.

SF-36 results for exit clients

As before, trial participants were grouped into three exit groups – those who remained in the trial, those who died or moved to residential care and those who left for some other reason. It was sensible to only look at the first and second applications because, by the time of the third application, the exits had almost all exited. When the exit groups were compared using data at each of these separate applications, overall significant differences were found. Moreover, there were significant differences amongst the three exit groups for PCS, MCS and all subscales when they were considered separately. Consistently, the group who remained in the trial achieved higher scores, while those who died or moved to residential care during the trial had the lowest scores.

Any improvement or deterioration in health between the beginning and the midterm of the trial should be reflected in the change in SF-36 scores during that period. Overall, a significant difference was found between the exit groups, but when PCS and MCS were investigated separately, it was found that the groups differed in the change in MCS scores only, and not in the change in PCS. There was little difference between those who stayed in and those who withdrew for some other reason; those who died or moved into residential care had much lower scores on MCS than the others.

Looking at the change in individual subscales between the initial and midterm applications, several statistically significant results were found. Significant differences were found between exit groups in the changes in Vitality and General Health. Again the pattern is that the no-exit group and those who exited for some other reason showed comparable results, while the group who died or moved to residential care declined much more quickly. Vitality declined significantly for all exit groups, especially the died/ residential care group. All exit groups improved significantly on the Mental Health subscale between the first and second applications, but all who left the trial improved less than those who stayed. The change in Role-Emotional also differed significantly between the exit groups. However on this subscale, the most improvement was in the other exit group, followed by those who remained. Those who died or moved to residential care declined considerably.

SF-36 results summary

- There was no statistically significant difference found between the intervention and the control groups in any of the three applications of the SF-36.
- No overall statistically significant differences were found between intervention and control groups when the difference in scores from one application of the SF-36 to the next were tested.
- There were statistically significant differences found between clients who died, those who went to residential care and those who completed the trial.

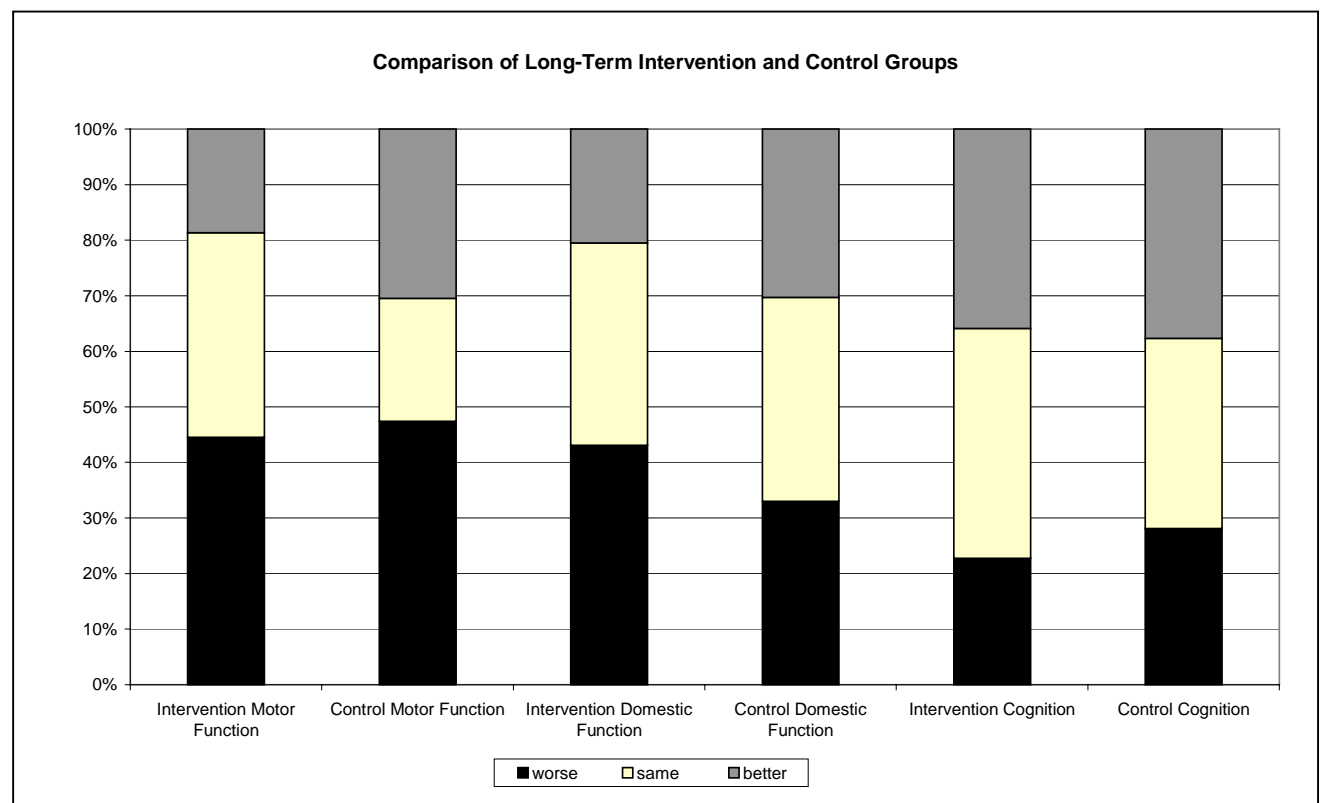
9.2.3.3 *Health and well-being as assessed through the Care Net assessment measures*

Three assessment measures were available on at least two occasions for both the intervention and the control groups. These were measures of motor function (FIM), domestic function (IADL) and cognitive function. For the purposes of comparing the long-term outcomes of the two groups, we analysed only records where there was at least a twelve month interval between the two scores on each of these instruments.

As seen in Figure 9, there were differences in the percentage of each cohort improving or declining on each measure. Both the intervention and the control groups showed most decline in motor function, with about 45% of both groups declining on this measure. While 30% of the

control group had a better score at their last assessment, only 19% of the intervention group did so. The least decline (and most improvement) was on the cognition scale. Again, this is partially attributed to the learning effect associated with using the same test on three or more occasions. About one third of both groups showed no change on the IADL measure. Of those who did change on this measure, about half of the control group improved while about half declined. For the intervention group, about one third improved while about two thirds declined.

Figure 9



Averages and standard deviations of the scores at each application of the tools are presented in Table 16.

Table 16 Assessment results for comparing long-term outcomes in the control and intervention groups

	Intervention (n=440)		Control (n=294)	
	Mean	Standard Deviation	Mean	Standard Deviation
First FIM score	86.9	6.95	87.7	4.89
Last FIM score	85.2	8.99	86.5	5.59
First cognitive score	3.0	4.17	3.1	4.0
Last cognitive score	2.5	4.40	2.9	4.5
First IADL score	12.0	5.13	11.0	4.5
Last IADL score	13.1	5.68	11.5	5.0

We were interested in whether the change in assessment scores over time was the same for both the intervention and the control groups. To answer this question, a multivariate analysis of variance (MANOVA) was used to compare the groups on the changes in all three assessment scores at once.

For all three assessment tools, the direction of the change was the same for both groups: if the average score decreased for the intervention group, then it also decreased for the control group. Average FIM scores were lower at the last assessment than at the first for both groups. This pattern was repeated in the cognitive function scores. However, average IADL scores increased from the first to the last assessments. Of course, higher FIM scores indicate better health while the reverse is true of the IADL and cognitive function scores.

Overall, the degree of change in health status as measured by these three assessment tools did not differ significantly between the control and intervention groups. Consideration of each assessment tool separately, however, disclosed that the increase in IADL scores was significantly less in the control group than in the intervention group. Even though the difference was statistically significant, it is doubtful that the difference was large enough to be clinically useful (average 13.1 versus 11.5).

In summary:

- There were no statistically significant differences between the intervention and control groups when all measures were taken together;
- There were no statistically significant differences between the intervention and control groups in terms of motor or cognitive functioning;
- There was a statistically significant difference between the intervention and control groups on the instrumental function scale with the control group doing better than the intervention group. But this result is not meaningful, as the difference in absolute values was very small.

9.3 How well did the Care Net assessment measures predict outcomes?

One of the potential uses of Care Net's assessment tools used is as a predictor of outcomes. We undertook a preliminary analysis of this potential use by splitting the intervention group into three outcome cohorts:

- Those who completed the two years of the trial;
- Those who exited the trial by moving to residential care; and
- Those who died during the trial.

The numbers of intervention clients in each cohort for whom data were available are shown in Table 17.

Table 17 Number of intervention clients in major exit cohorts

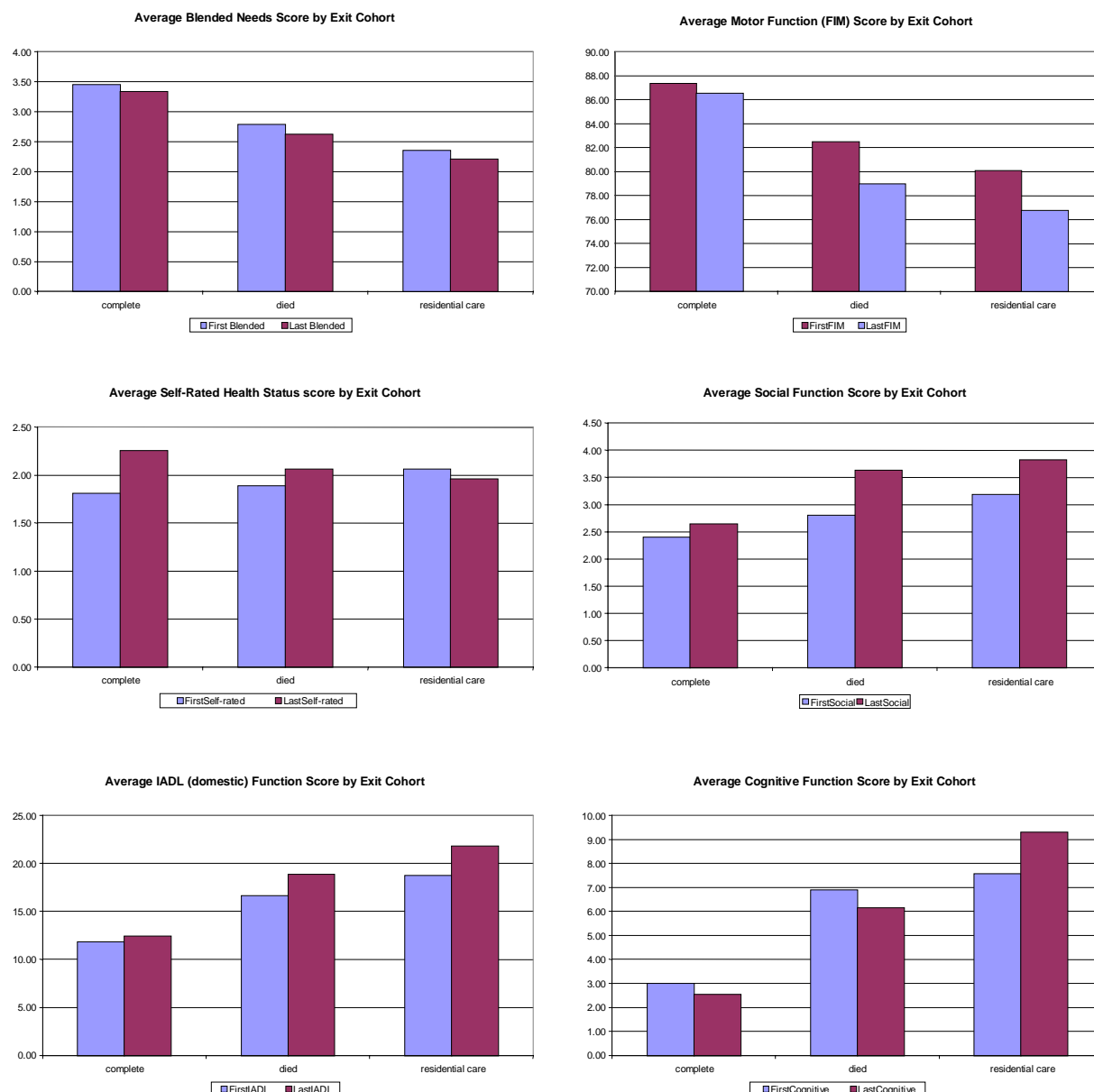
Exit Cohort	Blended	FIM	Self-rated	Social	IADL	Cognition
Complete	996	990	838	984	989	990
Died	98	79	63	79	79	79
Residential care	81	64	49	63	64	64
Total	1175	1133	950	1126	1132	1133

Figure 10 shows the average ratings at the first and the last assessment for the three outcome cohorts. Participants exiting the trial for other reasons (for example, because they moved to another area to live) were excluded from the analysis because of their small volumes.

Note that, for two measures (the Blended Needs Measure and the measure of motor function – the FIM) a lower score indicates worse health status, while the reverse is true for the other four measures. Taking this into account, the patterns shown in Figure 10 are quite consistent across the various measures.

People who exited the trial by moving to residential aged care had, on average, the worst scores at their first assessment and the worst scores at the time of the last assessment. Their ratings were worse than the people who died during the course of the trial. Intervention group clients who completed the trial had the best ratings at both their first and their last assessment.

Figure 10 The assessment tools as predictors of outcome



Since there were two sets of assessments for each client, the question arises as to how closely the two sets of scores are related. In Figure 10, it can be seen that the pattern of scores amongst the outcome cohorts is the same for the first and last assessments for five of the six measures. The Self-Rated Health Status score is the exception. At the time of the first assessment the group who completed the trial had the lowest average Self-Related Health Status score, while those who moved on to residential care scored the highest. However, in the last assessment the pattern was reversed.

This observed pattern between first and last assessment scores for each of the assessment tools is also reflected in their correlations. The two sets of Self-Related Health Status scores had a correlation of only 0.07. Correlations for the other measures ranged from 0.55 for the Social Function scores to 0.80 for the IADL (domestic) Function scores.

Another question of interest is whether any of these assessment measures can be used to assess the relative risk of falling into the different outcome categories. The appropriate statistical technique to answer this question is logistic regression. Using this technique, a statistical model is

built by testing the relevance of different explanatory variables in estimating the probability of belonging to a particular category of the response variable. Explanatory variables that are found to contribute significantly are left in the model. Thus it is possible to investigate the relationship between the assessment scores as explanatory variables and the outcome measure as the response variable.

As well as looking for statistically significant explanatory variables, there are several other ways of assessing how well the model fits the data. These include different ways of counting how often the actual outcome cohort of the trial participant coincided with that predicted to be most likely by the model. The odds ratio can also help in assessing the relationship between the explanatory variables and the outcome category. This figure gives the odds of falling into one outcome category relative to another. It can be interpreted as the multiplicative increase (or decrease if its value is less than one) in the likelihood of being in one outcome category when the explanatory variable increases by one unit. For some of the assessment tools, such as the FIM, a one unit change is too small to be of any interest, but a larger change, say 15 units, would be worth noting. For other tools, such as the Self-Rated Status score, there are only a few values possible, so a change in one unit is important. The odds ratio has to be interpreted bearing this in mind.

Each of the six assessment tools was tested individually. All except one, the Self-Related Health Status score, were found to have a significant relationship with the outcome cohort using the data from the first assessment. Using the last assessment data, all measures were found to be significant when they were tested one by one. However, the odds ratios ranged in value from 0.64 to a little more than 2. Thus the likelihood of being in a particular outcome category does not increase by a large amount if the explanatory variable increases by one unit. This is true for all assessment tools, those with many possible values as well as those with only a few.

An indication of how well the models discriminated between the outcome categories can be obtained by investigating the probability they assigned to the actual outcome. From this we can estimate how often clients would be assigned to the correct outcome group. Using the first sets of assessments, the models were found to be quite good, but the proportions correctly assigned to the actual outcome category were too low to justify the use of the models as predictive tools. Accuracy was higher in the models derived from the data from the last assessments. This is hardly surprising as these measures were taken closer to the time of the final outcome, so would be expected to relate more closely to it.

Not all measures were available for all clients. Although it meant sacrificing some data, a logistic regression analysis was performed fitting all variables together. Only clients with a full set of measurements could be included in this analysis. In this way, we can examine the relationship between the outcome categories and a combination of assessment tools and test to see if a subset of these measures can provide an adequate model.

When all assessment variables were entered in the model, just two were found to be statistically significant. These are the Cognitive Function and the IADL (domestic) Function tools. Adding other assessment tools did not significantly improve the model. With just these two explanatory variables, the model was reasonably accurate in estimating the likelihood of obtaining the outcome that actually occurred, but was not good enough to reliably predict the outcome category.

In summary, all assessment tools were found to be significantly related to the outcome when tested one at a time. When all variables were fitted together, only two were significant, the Cognitive Function and the IADL (domestic) Function tools; the other variables did not contribute significantly more, once these two variables had been entered into the model. However, the models provided insufficient discrimination to reliably predict outcome category.

9.4 The relationship between cost and need - how well did the Care Net assessment measures predict costs?

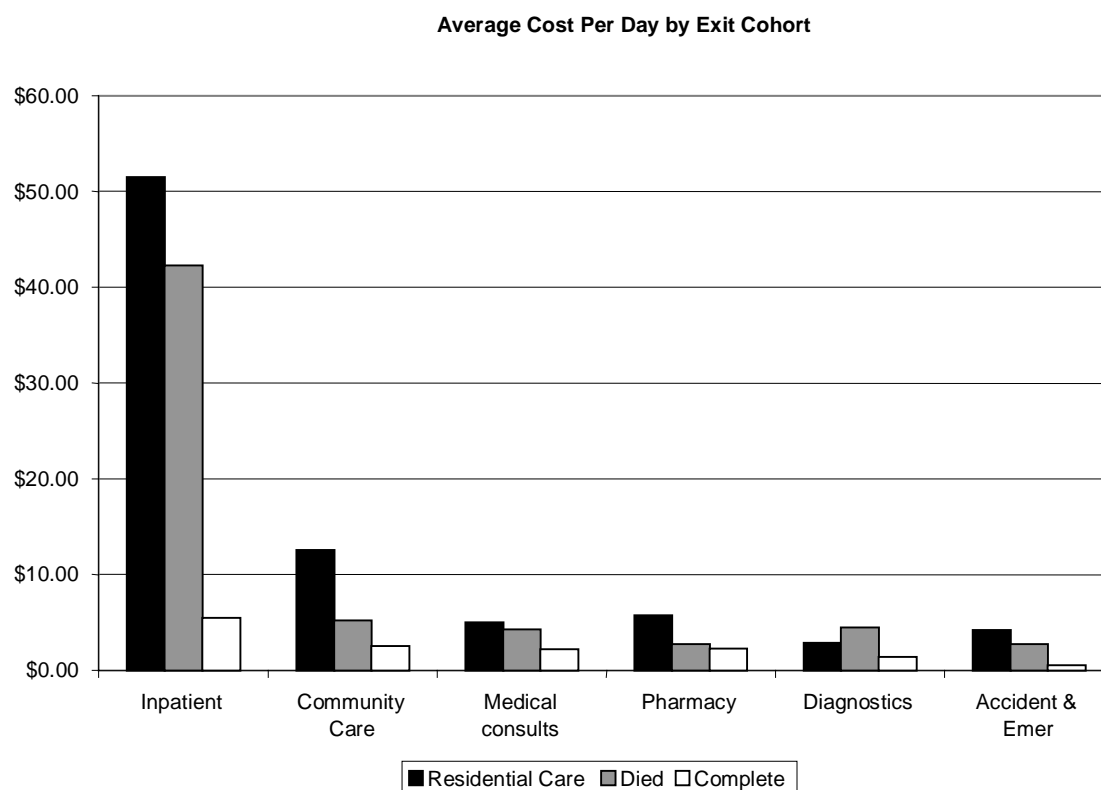
This section reports on the relationship between the intervention group’s need for care and the cost of the services actually provided. In all cases, costs have been calculated based on the actual number of days that each person was in the trial. The total costs have been divided by the actual number of days in the trial to derive an average cost per day per person.

9.4.1 The relationship between cost and exit outcome

Figure 11 shows the average daily costs for the three exit cohorts – people who died, people who exited the trial by moving to residential aged care and people who were still participating in the trial at the time of its completion. As before, clients exiting the trial for other reasons are excluded due to their small numbers.

Clients who went to residential care used more services on a daily basis than either clients who died or those who completed the trial. This is most notable for hospital inpatient and community care services. The only exception is for diagnostic services. For diagnostics, the highest users were those who subsequently died.

Figure 11



9.4.2 The relationship between cost and need as measured by the various assessment instruments

As described elsewhere, clients were assessed throughout the trial with a battery of assessment instruments. This section examines the relationship between the results of the first assessment and the subsequent costs of services consumed.

The overall results suggest that the some instruments used in the assessment process were significantly related to service costs. This complements the results reported previously that suggested that they were significantly related to subsequent outcome.

Figure 12 shows the cost of services based on the overall need level as measured at the first assessment. Intervention group clients classified as ‘high need’ on the blended needs index used the most services and incurred the highest per capita costs across all service categories. In turn, the ‘medium need’ group in turn incurred higher costs than the ‘low need’ group.

Figure 12

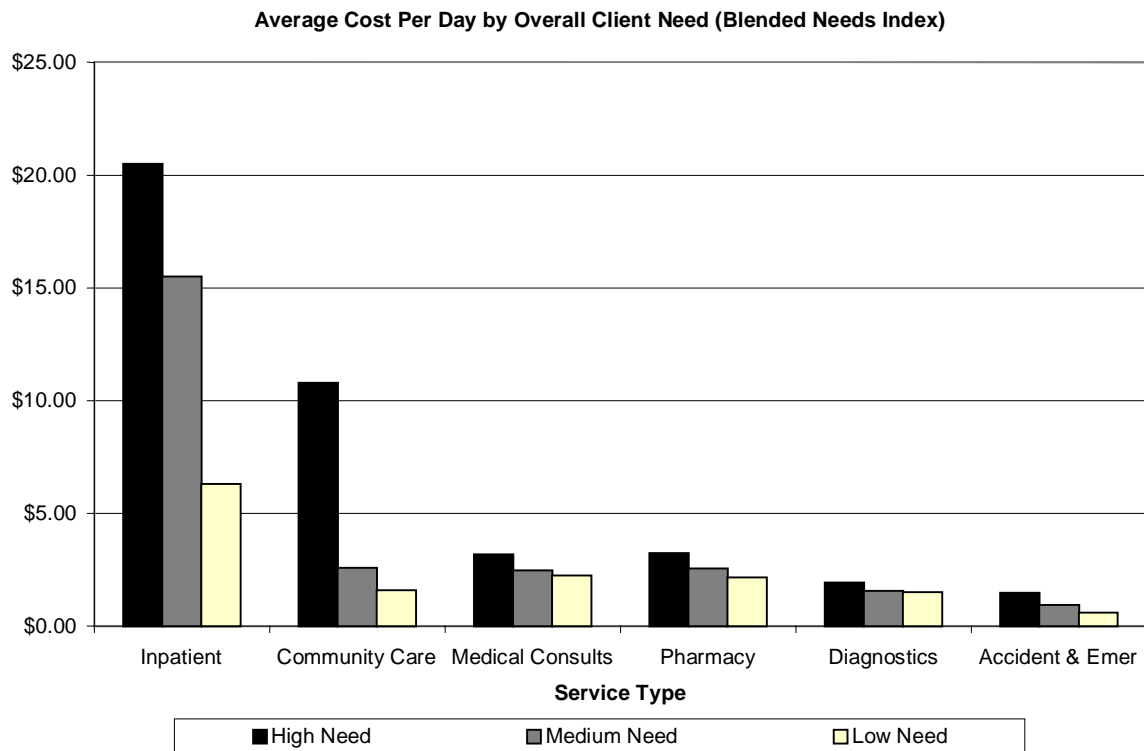


Figure 13 shows the same results but this time by examining the first assessment data on cognitive impairment. Clients with moderate to severe cognitive impairment used more services overall than those without cognitive impairment. However, most of this difference was driven by the difference in the use of inpatient services. There were differences between the service types. While those with cognitive impairment used more community care and accident and emergency services than those without cognitive impairment, they used less diagnostic services, pharmacy and medical services. This latter finding is surprising.

Figure 13

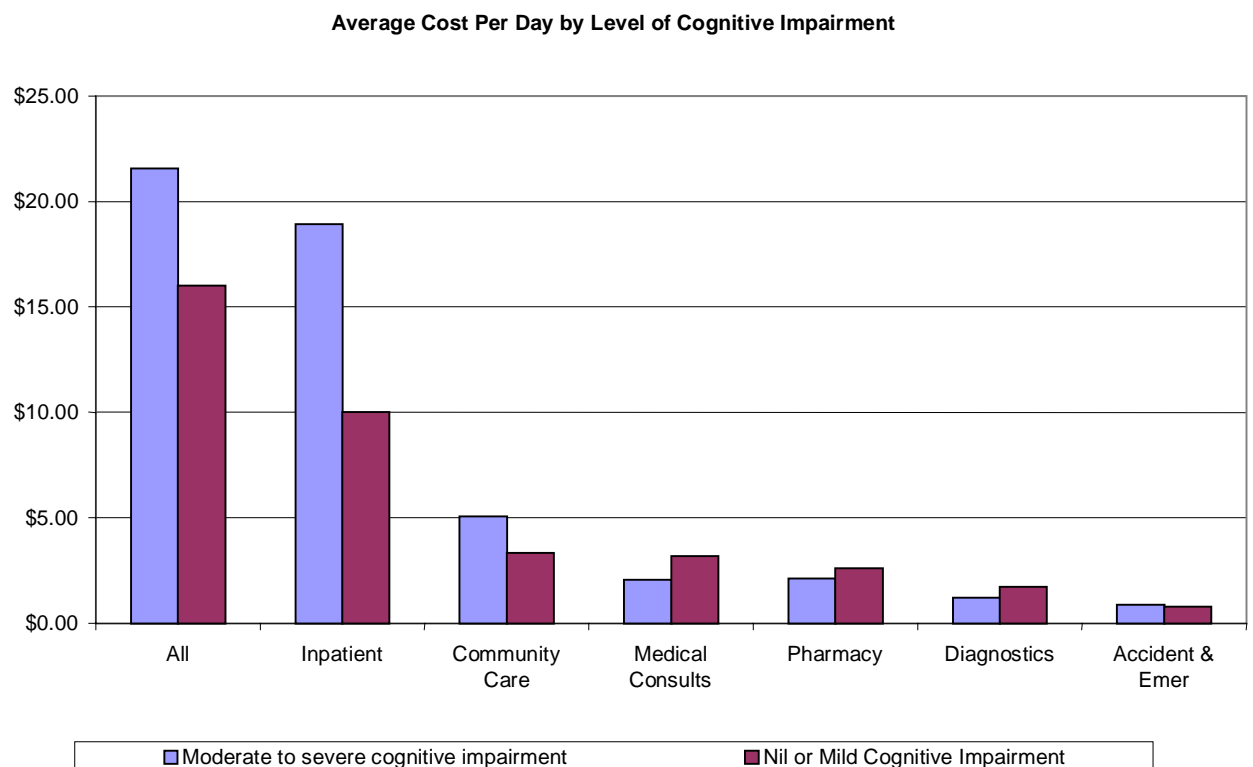


Figure 14 shows differences in the cost of services in relationship to instrumental (domestic) functioning. As can be seen, there was a logical relationship between IADL score at first assessment and subsequent costs, particularly the cost of inpatient and community care services. It also logically related to the costs of medical consultations, pharmacy and accident and emergency services, although the differences were not quite as marked. Instrumental functioning was not well related to the use of diagnostic services.

The results for social functioning shown in Figure 15 show a clear distinction between the two groups in relation to provision of community care. There were also differences in the costs of medical consultations, pharmacy, diagnostics and accident and emergency services although the differences were not as marked. Social functioning was not related to use of inpatient services.

Figure 14

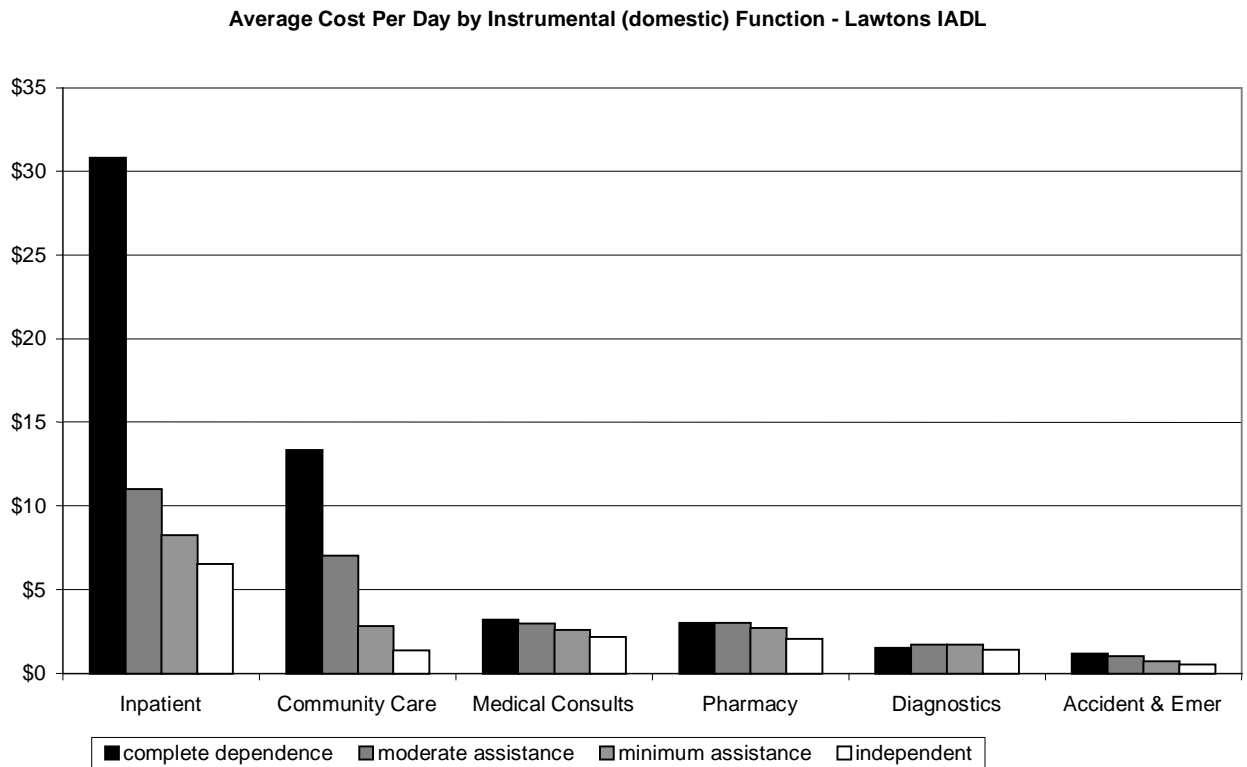


Figure 15

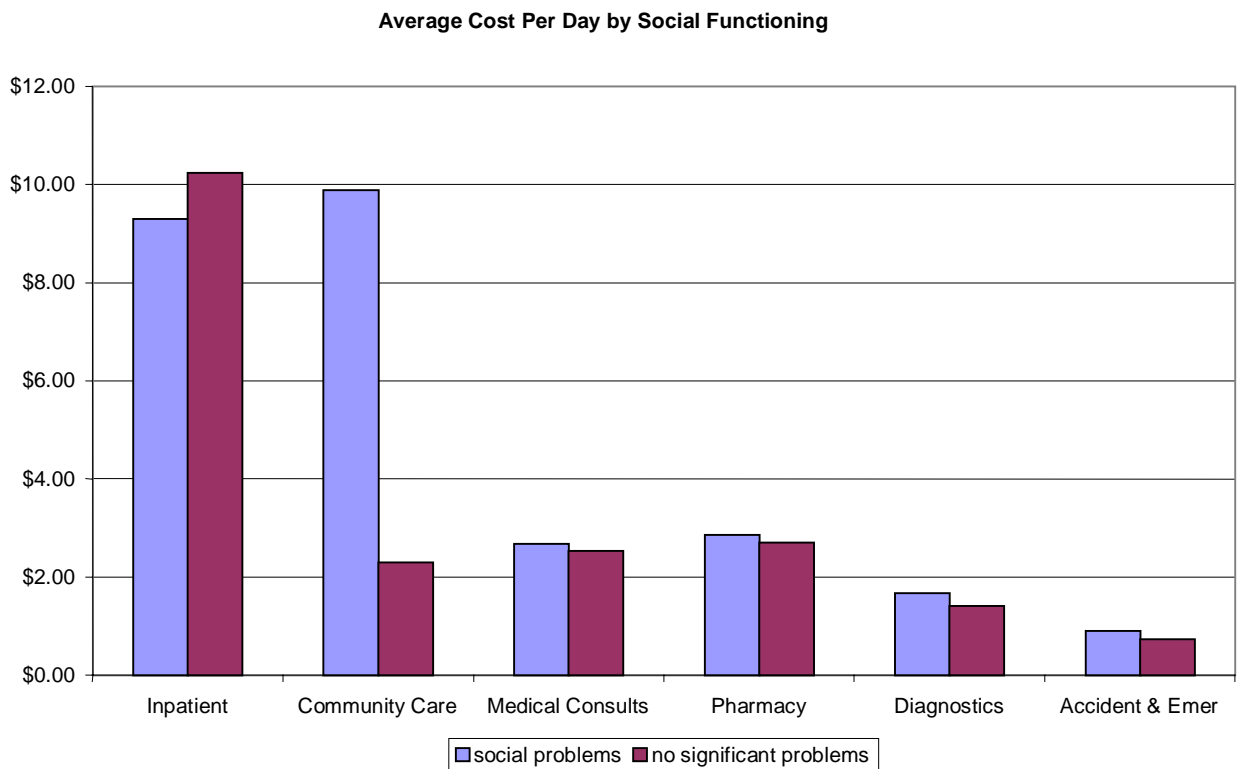


Figure 16

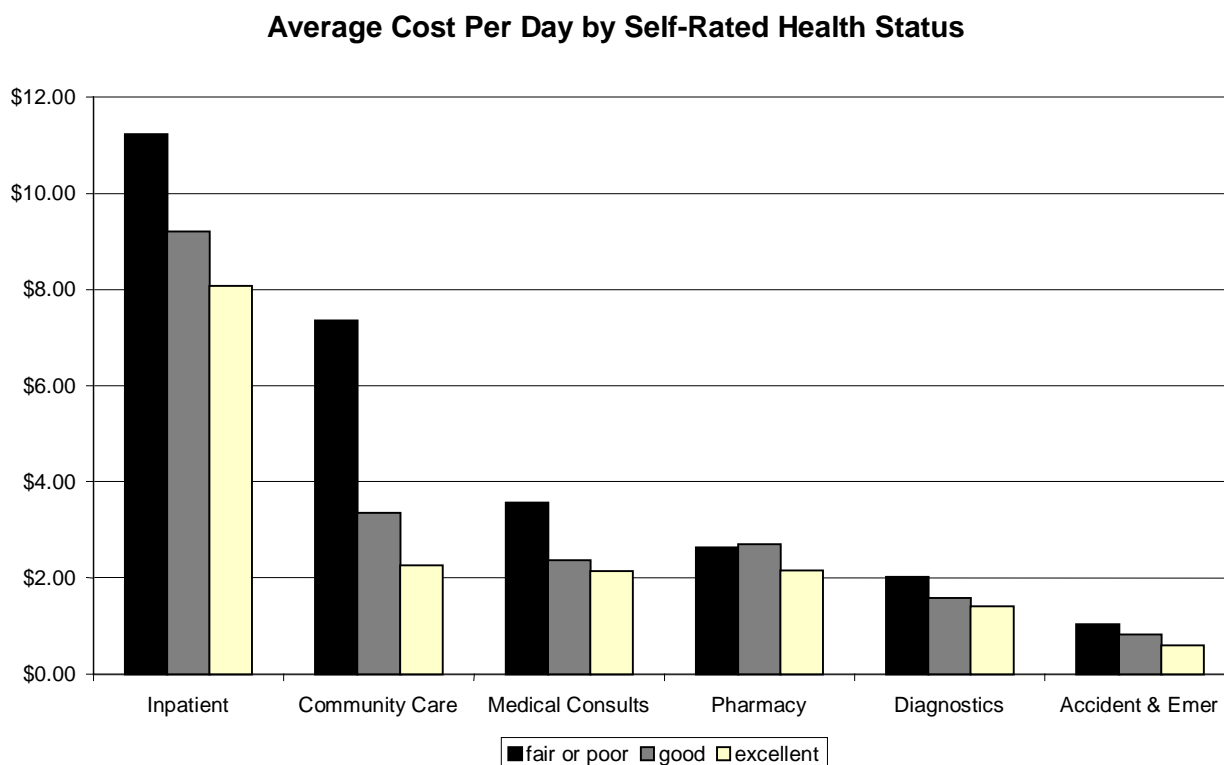


Figure 16 shows the relationship between the self-rated measure of health status and the cost of services. It was expected that those that rated their health as excellent or good would use less services than those who rated their health as fair or poor. This proved to be the case for all not diagnostic services. The most noticeable difference was in the provision of community care.

Motor function has been demonstrated in numerous studies to be highly correlated with the use of health and community care services. This was measured in Care Net by use of the FIM instrument. As expected, FIM score at first assessment proved to be related to the cost of inpatient and community care services during the trial. These results are illustrated in Figure 17. However, it will also be seen that the FIM was not related to the costs of the other services funded by the trial.

Figure 17

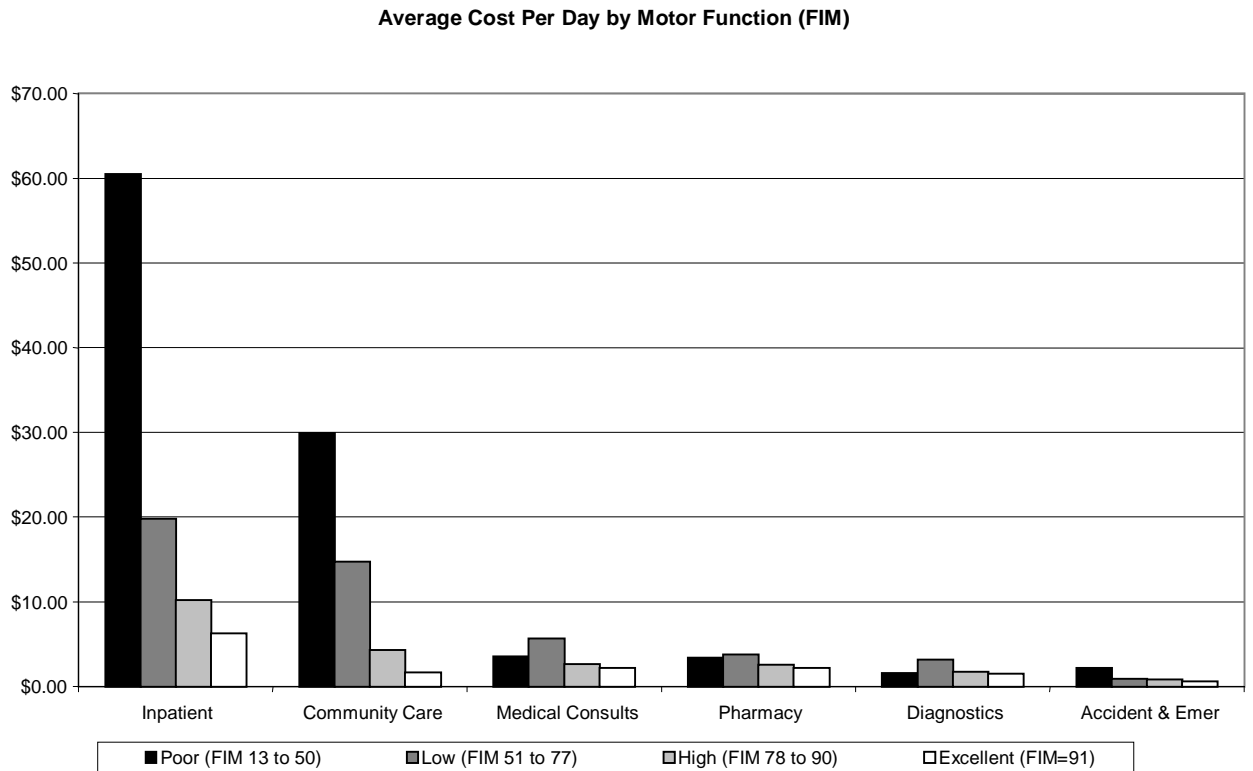


Figure 18

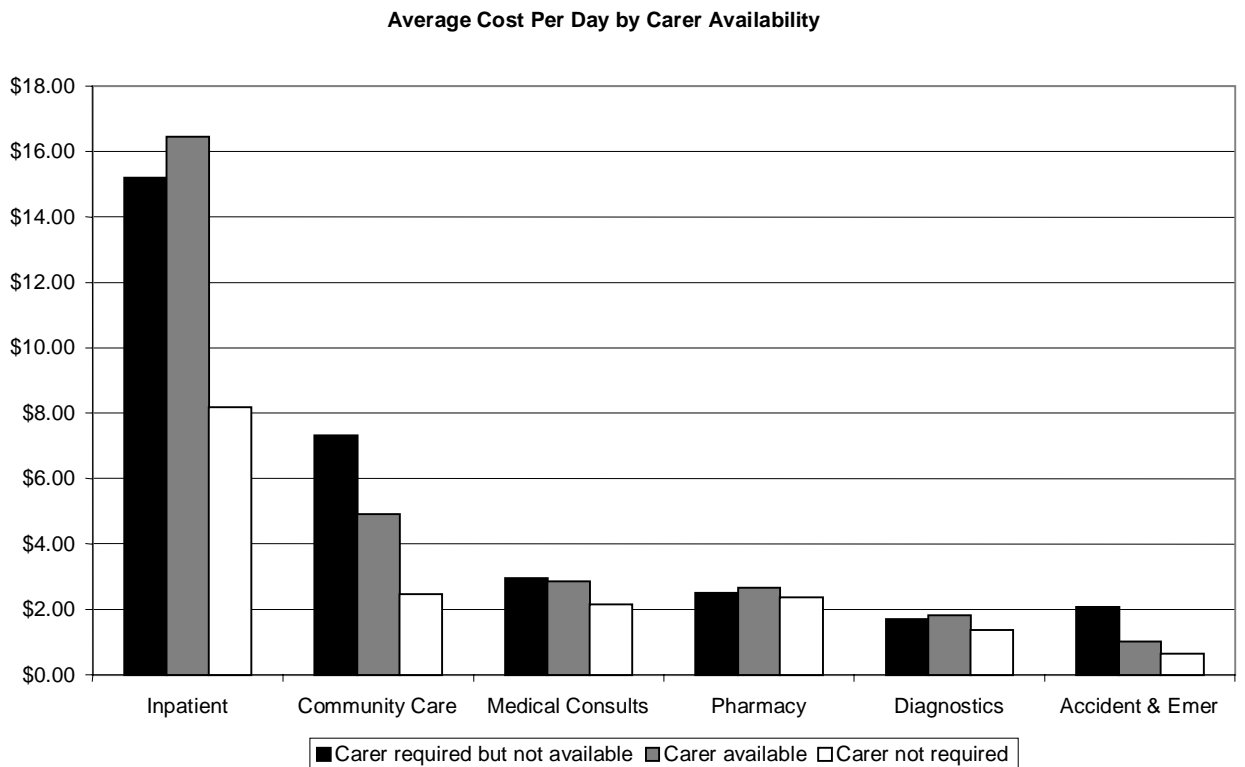


Figure 18 shows average cost per day by availability of a carer. It had been expected that those who were assessed as requiring a carer but not having one would consume the most services. Conversely, it was expected that those people not requiring a carer would use the least services. The results for community care, medical consultations and accident and emergency services conform with what had been expected. For inpatient services, those requiring a carer (who presumably were sicker) used more inpatient services than those not needing a carer. But whether or not a carer was actually available did not seem to be related to inpatient service costs.

9.4.2.1 Statistical testing of the relationship between assessment results and service costs

Multiple regression analysis was used to investigate the relationship between the measures described above and the average cost per day. Total cost per day was used as the response variable and different combinations of explanatory variables were used in different models. Exit status was used in the first model. Scores from the first and the last assessments were included as explanatory variables in two other models. Stepwise and backwards elimination techniques helped to determine a subset of statistically significant explanatory variables.

One other regression model was fitted using the change in assessment scores as the explanatory variable. For this model it was necessary to use the natural logarithm of the cost per day as the response variable. A small number of extremely low and extremely high costs were trimmed out as outliers. These were costs less than \$1 per day and greater than \$100 per day and constituted just over 3% of all values.

There are several ways of evaluating the regression analysis. The overall fit of the model is important. Models which fit the data well can be used for prediction and explanation of the relationships between the variables. One of the statistics which indicate how well a model fits is the R-squared value. Expressed as a percentage, it indicates the proportion of variability of the response variable that can be accounted for by the model. Explanatory variables that are found to be statistically significant are also of interest as they are making a useful contribution to the model.

None of the models was found to provide a good R-squared value. In fact not one was over 19%. This means that, although relationships amongst the variables can be investigated, it would be impossible to predict cost using the model. A lot more information would need to be added to the model to “explain” the average cost per day.

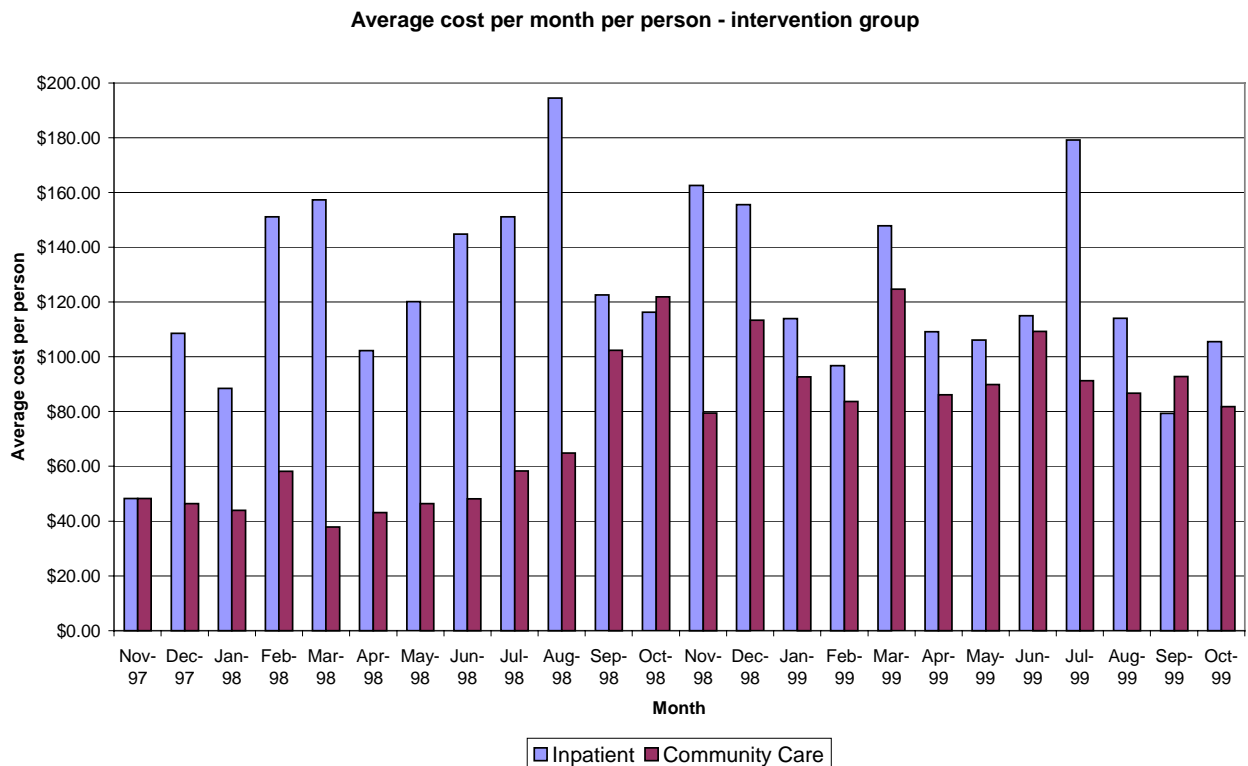
However, some variables were found to be significantly related to cost. Exit group was statistically significant, though the model had a very poor R-squared value. Using the set of first assessments, all but the IADL score were statistically significant. However, including just the two assessment tools, FIM and social function, produced a model which accounted for almost as much of the cost as the model including all assessment variables. Only the motor FIM and the IADL function scores were statistically significant in the model built with the set of final assessments. In the final model which fitted the difference in assessment scores, the IADL, the self-related health status and the blended needs scores were found to be significant.

When the exit status and assessment tools were fitted together as explanatory variables, a slight improvement in the explanatory power of the model was obtained. However, with R-squared values less than 32%, the models would still be inadequate for prediction of service costs.

9.4.3 Cost of services over time

Figure 19 shows the average monthly cost per person for inpatient and community care. Equivalent results for the other service categories are shown in Figure 20. In both cases, figures for the last two months of the trial have been excluded as not all invoices had been received at the time of preparing this analysis (March 2000). As with the previous analysis, the results are based on the actual number of clients in the trial each month.

Figure 19



There is some suggestion in the data that the cost of inpatient care began to decrease in the last six months of the trial. Some of this decline could be expected given the results shown in Figure 11. These results had indicated that people who went to residential care or who died were the more expensive clients. Those who completed the trial are over-represented in the results for the last six months because most of the people who died or went to residential care had already done so by the last stage of the trial.

There is also some suggestion that community care costs per person began to taper off after July 1999. While some of this may be attributed to the changing mix within the intervention group (see previous paragraph), it also reflects a decision by Care Net around that time to begin to reduce community services. For example, trial management took the decision not to approve home modifications in the last part of the trial. This decision was made in response to Care Net's worsening budget position and in preparation for the trial closure.

Figure 20

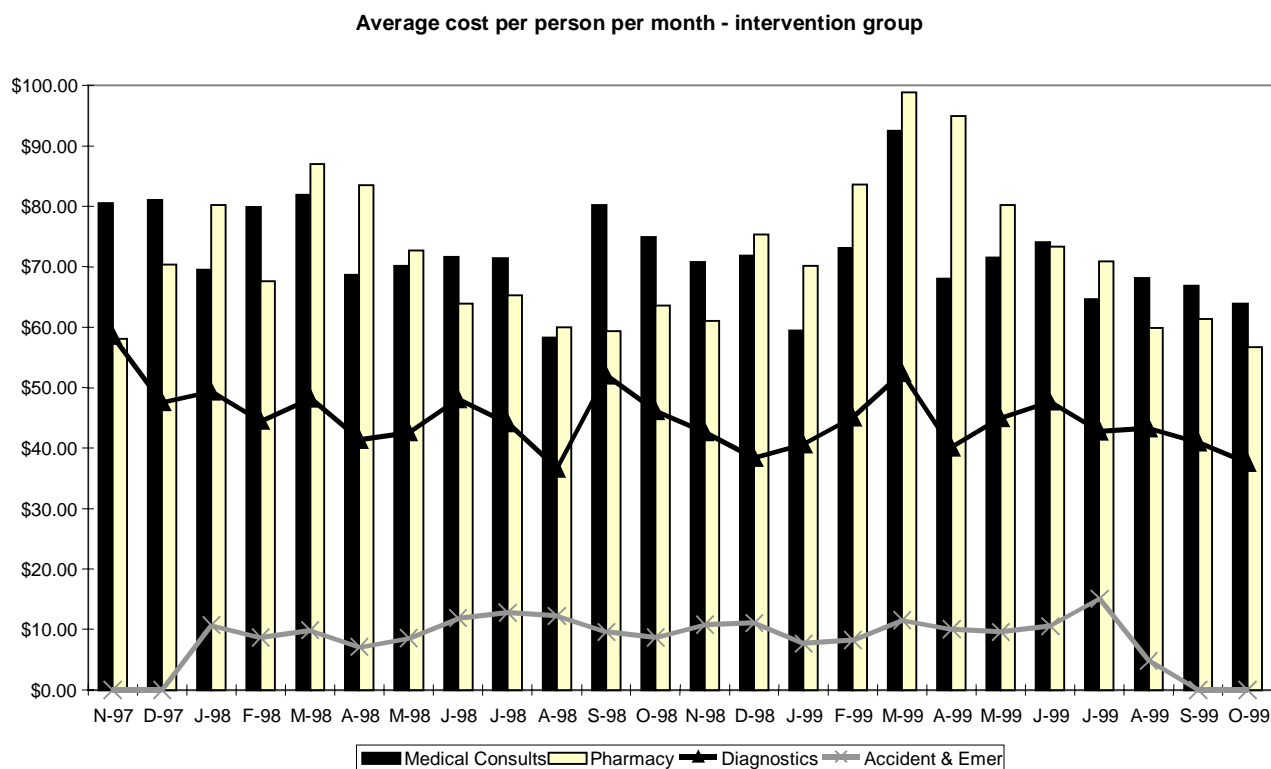


Figure 20 shows equivalent results for the other services. As before, there is some suggestion of a slight reduction in the costs of medical consultations and diagnostic services. However, the trend line for both pharmacy and accident and emergency services appears to be flat. There were no accident and emergency cost data available to the evaluators at the time of this analysis (March 2000) to cover the period before January 1998 or after August 1999.

9.5 Economic costs and evaluation

The financial results presented in earlier sections include only some of the costs incurred by participants. Intervention clients received a range of other services that were provided at no cost to Care Net and some economic costs were excluded from the Care Net fund pool. These costs, and their likely affect on Care Net's financial result, are summarised in Table 18.

Table 18 Costs excluded from the Care Net bottom-line

Excluded cost	Likely impact
Private hospital and auxiliary services	Increase in service utilisation costs (18% of participants had health insurance at the beginning of the trial and a further 11% had DVA cards). Impact of including private hospital utilisation on bottom-line result is unclear. External changes during the course of the trial (extension of DVA 'gold card' status and introduction of financial incentives to take out private health insurance) could have increased private hospital use and, conversely, reduced use of IAHS services.
Inter-regional flow for public health care	All services provided outside IAHS were excluded. Inclusion would have increased service utilisation costs. If included in Care Net fund pool, would have increased both income and expenditure, probably with net effect.
Non-PBS funded pharmaceuticals and other consumer out-of-pocket expenses.	Inclusion would have increased service utilisation costs. If included in Care Net fund pool, would have increased both income and expenditure, probably with net effect.
HACC services other than Home Care	Inclusion would have increased Care Net fund pool. However, as current clients continued to receive pre-trial levels of service at no cost to Care Net, the net effect is likely to be marginal.
IAHS excluded services (acute outpatients, mental health, intensive care, dental services)	Inclusion would have increased service utilisation costs. If included in Care Net fund pool, would have increased both income and expenditure, probably with net effect.
Information technology capital costs (\$1 million)	Depreciation of IT would have increased expenditure and increased Care Net bottom-line deficit.
Costs meet by IAHS and not charged to Care Net (software licenses and related)	Inclusion would have increased expenditure, increased Care Net bottom-line deficit and reduced IAHS liabilities.
Costs carried by participating agencies, particularly in data management and billing systems	Inclusion would have increased expenditure and increased Care Net bottom-line deficit.

A full economic costing and evaluation would allow for these items to be incorporated into Care Net's final results. The question arises of whether a full economic costing and evaluation is warranted.

In our view, a full economic evaluation is not warranted. A full economic evaluation would only be warranted if three conditions were met:

1. If Care Net clients liked the services that they received; and
2. If the outcomes for intervention clients were at least as good as those for control group clients; and
3. If there was any likelihood that it would change the bottom-line.

The first condition has been met. As described in Report 5 in this series, Care Net clients were overwhelming positive about their Care Net experience.

The second condition has been partially met. Overall outcomes for consumers were no different between the intervention and control groups. On most measures, they were the same. However, on one measure – rate of admission to residential care – the control group did significantly better

than the intervention group. There is also some (weak) evidence that the control group did better than the intervention group in relation to their ability to perform domestic acts of daily living.

However, the final condition has not been met. Even if all of the unreported costs were measured and included, there is no reasonable likelihood that their net effect would change the bottom-line. Care Net met the perceived needs of its clients and, not surprisingly, the clients were overwhelmingly positive in response. But, after two years, there is no evidence that outcomes for consumers actually improved as a result of the Care Net intervention or that the Care Net model is sustainable within current resources, however they are measured.

A full economic evaluation would also have measured the 'opportunity cost' of the Care Net model. The additional funds incurred on Care Net clients were not available for the use of other (high need) consumers and for use in the delivery of other (potentially more sustainable) models of care. Opportunity costs in terms of both population equity and sustainable service delivery are impossible to measure.

Nevertheless, Care Net achieved a critically important outcome, the results of which are too early to measure. It provided important lessons to the broader health and community care sectors on what works and (equally important) what does not, and positioned the Illawarra to develop more sustainable models in the future. If the lessons from Care Net are learned and applied, the benefits may outweigh the costs.

9.6 Care Net – results relative to other trials

We do not have access to the raw data collected in the other trials. However, our impression, based on our discussions with other local evaluators, is that Care Net collected more systematic and comprehensive data than most of the parallel trials. In particular, the Care Net assessment data were considerably more sophisticated than most other trials.

This results in an interesting dilemma. As local evaluators, we had access to data that were not available either to the national evaluators or to other local evaluators. The dilemma is this: given that we had access to better data, have we raised the bar too high by assessing Care Net using a different set of standards than those applied to the other trials?

We do not think so. Care Net was a trial, not a new service delivery system. It was designed to test a series of hypotheses and the data that it collected were appropriate for that purpose. Those data demonstrate that the central hypothesis (*that coordination of care of people with multiple service needs, where care is accessed through individual care plans, and funds pooled from existing Commonwealth, State and joint programs, will result in improved client health and well being within existing resources*) was not supported by the evidence.

Appendix 1: An overview of the Illawarra Coordinated Care Trial

The Illawarra Coordinated Care Trial, or Care Net Illawarra, was established as part of a National series of demonstration projects that aimed to assess the benefits of coordinated care in the context of the Australian health care system. The primary hypothesis to be tested by the trials was that:

That coordination of care of people with multiple service needs, where care is accessed through individual care plans and funds pooled with existing Commonwealth, State and joint programs, will result in improved client health and well-being within existing resources.

Fund pooling was a key component of the coordinated care trials. The established, program-based organisation and funding of health care services was thought to be a principal factor in frustrating flexible service provision and service substitution. The way to overcome this issue was believed to be the pooling of funds from Commonwealth, State and joint Commonwealth-State programs (Pekarsky, 1999). Trials would be allocated a budget from which they could purchase services for clients, with funds for the budget coming from the finances of existing service providers. But there would be no additional money. A key Commonwealth requirement was for trials to be cost-neutral.

The Illawarra trial encompassed three local government areas: Wollongong, Shellharbour and Kiama, the area being located south of Sydney, NSW. Its principal stakeholders were the Illawarra Area Health Service (IAHS), the Illawarra Division of General Practice (IDGP) and the NSW Home Care service.

The trial aimed to coordinate the care of people aged 65 years and over with either a risk of falling or who had complex medical or social problems that required multiple services from more than one health care service provider. During the 1997 planning phase of the trial, roughly 1800 eligible residents were referred to the trial by the 100 GPs participating. 1200 clients were allocated to an active group and would have their care coordinated by the trial, while the other 600 were allocated to a control group.

After the planning phase, the trial went live on 1 November 1997. It finished on 31 December 1999, a total period of 26 months over 3 financial years.

The coordination of care was undertaken by 15-16 care coordinators, in collaboration with the client's GP. The GP maintained control of the medical aspects of the client's treatment, while the care coordinator organised access to other services, purchasing services agreed to with the participant and GP. The care coordinators performed a systematic assessment of their clients initially every three months and subsequently in response to need. These assessments were intended to inform the creation of the clients' care plan that, among other things, included the goals of care for the client. Finally, a service plan was created that described the package of services to be bought by the trial in order to address the clients' goals was created.

Thus, the trial adopted a commissioning model of service provision. It secured services mainly from community care service providers, having access to both public services (those in the IAHS and the local HACC agencies) and private services. Medical (GP, specialist), pharmaceutical and hospital services were also within the funding pool, but these services were largely determined by GPs.

The trial used State and Commonwealth funds to develop an information system. Based on an intranet structure, the system supported email, client records, service utilisation, and financial data. This system did not support full communication between providers but significant steps included the increased use of computers by GPs.

References

- Cromwell, D., Eagar, K. and Owen, A. (1997) *An estimate of the size of the Funding Pool for the Illawarra Coordinated Care Trial*, Centre for Health Services Development, University of Wollongong.
- Cromwell, D. (1998) *Estimated HIC pool contribution for PBS Services ("Live phase method" data)*, Centre for Health Services Development, University of Wollongong.
- Cromwell, D. (1999) *Second revision of the financial contribution to the Illawarra Coordinated Care Trial by the Illawarra Area Health Service*, Centre for Health Services Development, University of Wollongong.
- Cromwell, D., Eagar, K. and Owen, A. (1999) Filling a pool: estimating the cost of care for the Illawarra Care Net Coordinated Care Trial in *The Australian Coordinated Care Trials: Methodological issues in trial design and evaluation*, Vol. 2. Commonwealth Department of Health and Aged Care, (Eds, Leigh, J.) Canberra.
- Eagar, K., Owen, A., Perkins, D., Quinsey, K. and Adamson, L. (1998) *Care Net Illawarra: The Mid Term Report*, Centre for Health Service Development, University of Wollongong.
- Pekarsky, B. (1999) Service substitution: evaluation issues in the coordinated care trials In *The Australian Coordinated Care Trial: Methodological issues in trial design and evaluation* (Eds, Leigh, J.) Commonwealth Production Unit, Commonwealth Department of Health and Aged Care, Canberra.