FACTORS INFLUENCING THE OUTCOME OF COMMUNITY CARE IN A QUICK RESPONSE TRIAL IN ST. JOHN'S NEWFOUNDLAND

GWYNEDD BARROWMAN

FACTORS INFLUENCING THE OUTCOME OF COMMUNITY CARE IN A QUICK RESPONSE TRIAL IN ST. JOHN'S, NEWFOUNDLAND

by

Gwynedd Barrowman

A thesis submitted to the

School of Graduate Studies

in partial fulfilment of the

requirements for the degree of

Master of Science

Faculty of Medicine

Memorial University of Newfoundland

1998

St. John's Newfoundland

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

0-612-34162-3

Canada

Abstract

Quick Response (QR) Programs have been implemented in many places to assist patients and their families when through ill-health they are suddenly unable to manage at home. The aim is to tide them over a short-term crisis by providing speedy practical support in the home, to avoid unnecessary hospitalization, and where appropriate to provide information to assist families in making decisions for the future.

This work summarizes the achievements of a number of QR programs in Canada, and outlines the methods and results of the QR project carried out in St. John's, Newfoundland in 1995. This was a clinical trial in which appropriate patients who came to the Emergency Room (ER) were randomized to receive either QR in their homes, or standard care, in hospital or at home, as ordered by the physician. Health status and function in daily living were measured using three well-validated questionnaires: the modified Barthel, the Short Form-36 and the Sickness Impact Profile. Care provided over three months, and the associated costs, were documented. Caregiver stress was measured using the Relatives' Stress Scale and the General Health Questionnaire-30. Enrolment in the study was poor: possible explanations are presented.

Care in the community will play an increasingly important role as the population of Canada ages. The objectives of a community health agency are not the same as those of an acute care system; in the context of community care for people referred from the ER, some definitions of success are proposed. Based on data from the QR project,

increasing age and the presence of one or more adverse diagnoses such as congestive heart failure, chronic obstructive lung disease and dementia were identified as predictors of an unsuccessful intervention by Community Health, whereas a supportive network of family and friends increased the likelihood of success. Finally, some suggestions for future research and development in community health care are presented.

TABLE OF CONTENTS

List of Tables

List of Figures

Acknowledgements

I wish to acknowledge with gratitude the guidance, teaching and support given me by my supervisor Dr. Brendan Barrett. Thanks are due to the other members of my supervisory committee, Dr. V. Gadag and Dr. J. Segovia for their constructive criticism and advice, and to Gregory Doyle, Jackie McDonald, Sylvia Ficken and Sebastian Higgins for their help. I also thank the staff of Community Health - St. John's Region, in particular the nurses who took part in the QR project.

Finally I thank my family and friends for their constant encouragement.

Chapter 1 - Introduction

Quick Response Programs have been set up in a number of centres in Canada, to serve patients who are suddenly unable to manage at home. Injury or worsening health of a vulnerable person or of a suppporting family member may make a previously manageable arrangement unworkable, and unless additional help is provided quickly, hospital admission may be necessary. The swift provision of short-term homemaking, nursing and other professional services, at a level greater level than that usually provided, may help these patients and their families deal with the situation, and thus avoid unnecessary hospitalization. There has been little objective evaluation of these Quick Response programs.

The Quick Response (QR) project in St. John's was a research-oriented initiative to develop, operate and evaluate a QR program. Evaluation of the program was in the form of a clinical trial, with patients who agreed to take part being assigned randomly to receive either routine care, as ordered by ER and consultant physicians, or immediate QR services. This account describes the rationale, method and major results of the QR project.

My original intention was to use the data collected to identify factors which predict success in a QR intervention, but with low enrolment in the project, too few subjects for this purpose were assigned to receive QR. The focus of my study was therefore broadened to include all the subjects enrolled in this study, to identify some of the factors which predict success or failure in a wider range of Community Health interventions.

1.1 Health care in St. John's: The subject of this study is health care delivery in the home, in an urban community, St. John's Newfoundland. To appreciate the role of Home Care (now known as Community Health) it is worthwhile considering briefly the history of health care in St. John's.

Early records of health care delivery systems in Newfoundland are sparse, and it was not until the 1930s that organized records were maintained. The first resident physician arrived in St. John's in 1784 as medical officer to the garrison. By the start of the nineteenth century, there were two small military hospitals, and in 1814 a hospital for civilians was established. An institution for the care of psychiatric patients was opened in 1845 on the site of the present Waterford Hospital. In 1870, the military garrison left St. John's, and the civilian hospital took over the Forest Rd. site which is now the home of the Leonard Miller Centre.

Well before this, in 1800, Dr. John Clinch had introduced vaccination with cowpox to the residents of Trinity and then St. John's where "the smallpox was making great ravages"[1]. The first formal provision of health care in the community was made under the general heading of Public Charities, with the appointment in 1826 of a salaried public medical officer. Living conditions in St. John's were poor; the streets were filthy and sewage ran in the gutters. Epidemics occurred frequently[2] and in 1833 provisions were made for the introduction of quarantine measures where appropriate to limit the spread of infectious and contagious diseases[3]. In 1847, Lieutenant-Governor Le Marchant recommended construction of a sewage system, but shortage of funds delayed implementation for many years. After the 1854 cholera epidemic in the city in which 500 people died, an Act of Assembly (1857) provided for collection of statistical

information, and for the appointment of three more district medical officers.

Although a number of English nurses went with Grenfell to Labrador in 1893, it was not until 1895 that the first trained nurse was appointed to the General Hospital in St. John's. Tuberculosis was rife at the beginning of the century, and in 1909, the Association for the Prevention of Tuberculosis was formed. In 1911 summer camps started for TB patients, in tents set up in the grounds of the General Hospital. This was followed in 1917 by the opening of a 52-bed sanatorium, and in the same year, the formation of the Child Welfare Association. In 1922 and 1923, St. Clare's Mercy Hospital and the Salvation Army General Hospital respectively were founded. In 1929, in the context of extreme poverty, the Newfoundland Board of Health was established to supervise public health measures. Its main focus was the control of infectious diseases (notably TB, diphtheria, whooping cough, measles, typhoid and venereal disease). Also Included in the public health mandate were food inspection, maternal and child welfare, health education and improvement in nutritional standards.

In 1948, the year before Newfoundland joined with Canada, the death rate from tuberculosis in St. John's was 102 per 100,000, much improved from 190.6 in 1937, but still a major problem*. Infectious diseases continued to dominate the scene, with poliomyelitis epidemics in 1953 - 1954 (176 cases with 11 deaths) and 1959 - 1960 (188 cases with 17 deaths).

* By contrast, the Acquired Immune Deficiency Syndrome (AIDS) epidemic, while of great concern, seems unlikely to reach comparable proportions in this province. Since 1984, when data for the disease in Newfoundland were first collected, the cumulative provincial total as of September 1997 is: HIV positive, 178 cases; full-blown AIDS, 64 cases with 52 recorded deaths (information courtesy of Dr. Faith Stratton, Community Health.)

Rarely was anyone admitted to hospital with a stroke or the incapacity of old age: rather they were cared for at home, as best the family could manage, with occasional visits from a family doctor if they could afford to summon him.

With improvement in socioeconomic conditions, and the introduction of antibiotics (particularly streptomycin in 1947 and isoniazid, another antituberculous agent, in 1951) patterns of medical care changed. Tuberculosis, once the scourge of Newfoundland, receded into the past; the sanatorium, by then known as the Hospital for Chest Diseases, was finally closed in 1973, with the few remaining patients transferred to St. Clare's Mercy Hospital. Infectious diseases appeared to pose little or no threat, and other diseases such as stroke and heart disease began to attract more attention. The emphasis on treating the sick in hospital increased in the 1970s, and although Public Health and VON continued their valuable work in the community, their profile was low.

Now in the last years of the twentieth century, people are living longer: at birth, a boy born in Newfoundland in 1971 could expect to reach 69.3 years, and a girl 75.7 years. By 1996, baby boys had a life expectancy of 75.1 and girls, 81.8 years; a continued slight increase in life expectancy is projected at least until the year 2011. With migration of many young people from the province, we are suddenly aware that the population is aging dramatically: in 1980, the provincial population of individuals aged 65 years and over was ~42,600; by the year 2011, this figure is projected to double. In the 1794 census, the population of St. John's and the neighboring area was recorded as - 4,000, growing in 1891 to 29,000 (about 16% of the population of the province), and in the 45 years since then it has increased more than two and a half times, (Census Metropolitan Area, 1951 - 1996, Figure 1.1),

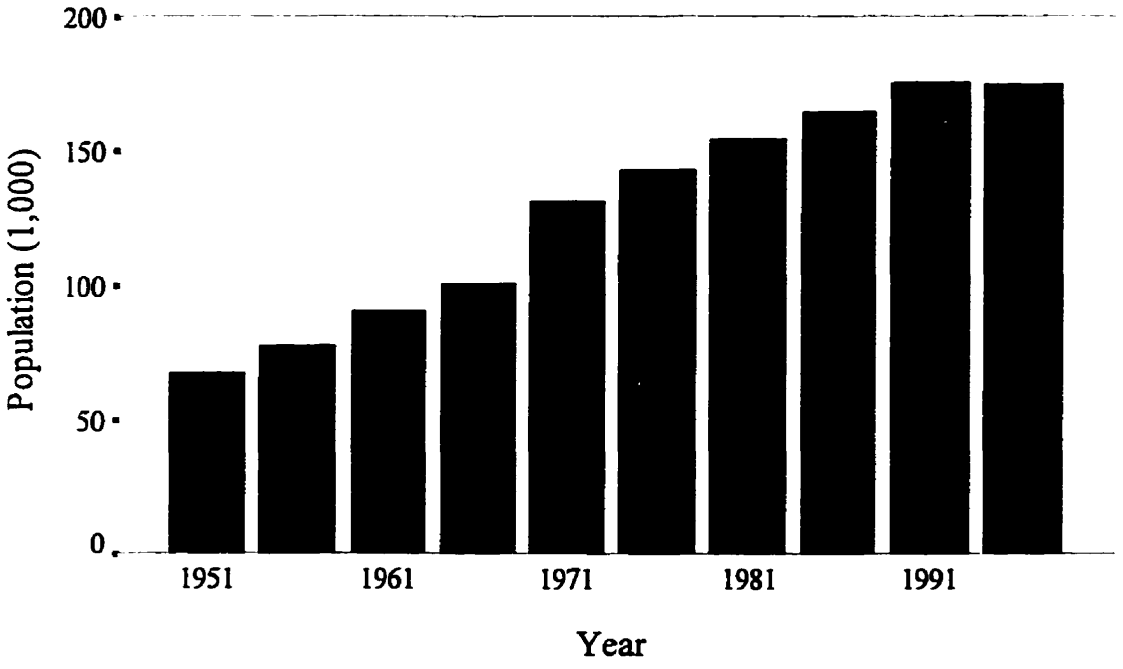


Figure 1.1 Population of St. John's, Newfoundland

St. John's Census Metropolitan Area

Statistics Canada, Demography Division, courtesy of Newfoundland Statistics Agency

while the provincial population has increased by a factor of only 1.5. There has been an inexorable trend for people to move from the outports into larger centres, and the population of St. John's as a percentage of the provincial population has increased from 18.7% in 1951 to ~ 30% in 1996. Moreover, the percentage of St. John's residents aged 65 years and over has risen, from 8.1% in 1981 to 10.7% in 1996. Part of this

increase is due to aging of longtime St. John's residents, and part reflects the movement of older folk into the city. At the same time, many younger families have moved away in search of work: the population of people aged 24 years and less has fallen by 15% in the fifteen years leading up to 1996. The concept of families caring for their elderly relatives and neighbours has always been a strong feature of Newfoundland society, but the emigration of so many young people is making this more difficult.

1.2 St. John's Home Care* (SJHC) was initially set up as a pilot project in 1974 to facilitate early discharge from hospital. The program clearly met a need and grew rapidly. Most patients received nursing care, but physiotherapy and occupational therapy, social work and laboratory services were also available. In the financial year 1977 - 1978, 774 patients were registered, and total expenditures reached over $298,000. The principal aims of SJHC were to free up hospital beds by preventing or shortening hospital admission; to co-ordinate existing services provided at home under the care of the patient's physician; and to provide continuity of treatment and rehabilitation. In 1981, the COLD program, designed for patients with severe chronic lung disease, was put in place and in November 1982, the Home Support program was added. In 1982 - 83, 1724 new patients were registered, with total expenditures of $836,457. The importance of longer-term support (provided to 12% of new patients in 1984 - 85) was increasing, and in 1985, the Continuing Care program was initiated.

* Information extracted from St. John's Home Care and Community Health (St. John's Region) annual reports, courtesy of Ann Crowley, Continuing Care Manager.

The area covered by SJHC expanded beyond city boundaries, and in 1986, SJHC assumed responsibility for Continuing Care services previously provided by St. John's Public Health Unit; the annual report for 1986 - 87 notes 3466 new patients and expenses of $1,742,000. By 1989, the workload had increased so much that clients with adequate insurance coverage were directed to private agencies such as the Victorian Order of Nurses (VON). Expenditures continued to rise, surpassing $3,900,000 in 1992 - 93. Of the 3007 new patients registered that year, 35% received long-term care. The biggest change to date came in April 1994 when SJHC, St. John's and District Health Unit and St. John's Drug and Dependency Services merged to form Community Health (St. John's Region). There were wide-ranging organizational changes; new initiatives such as single-entry to Continuing Care (community and institution-based) were implemented in 1995, and satellite offices were opened so that staff are based nearer the communities they serve.

Financial constraints have forced a re-evaluation of the delivery of health care; economic realities demand a more thorough assessment of efficacy and cost-benefits. It has become clear that while certain health problems are best dealt with in hospital, many others can be resolved or looked after with care delivered in the home, in day hospitals, or in long-term institutions. One person staying in hospital for one day was estimated to cost the public purse over $500; in a chronic care institution, $150, and in a personal care home, $29.* The Royal Commission which reported on hospital and nursing home costs in 1984 made about 200 recommendations to reform and rationalize the provincial health system. A major consequence of this report has been a

* 1994 - 1995 figures, courtesy of Dr. D. May, Department of Economics, Memorial University.

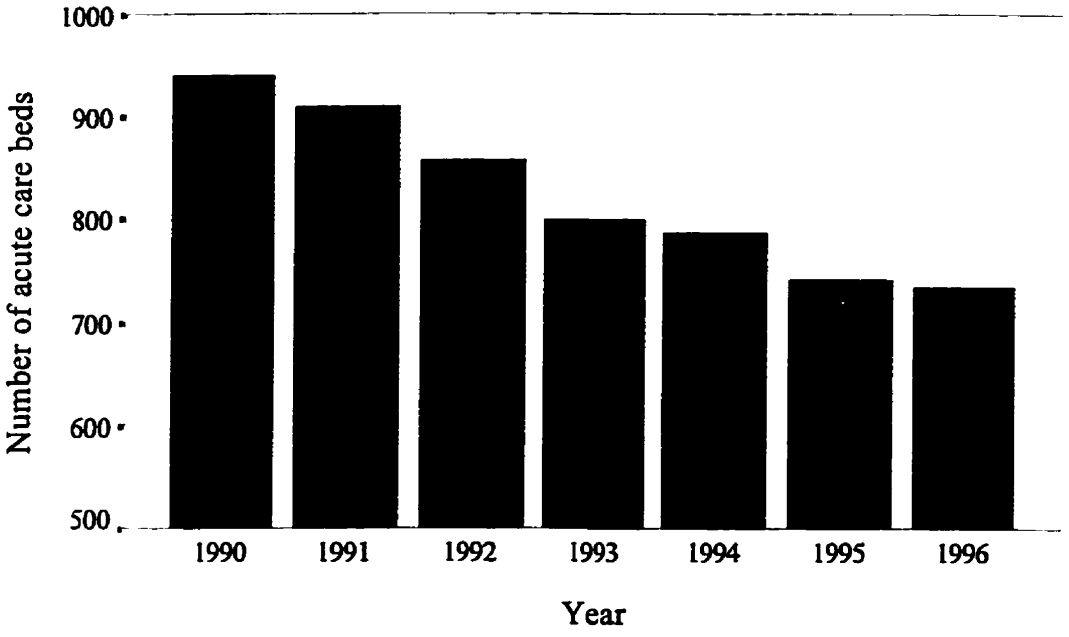


Figure 1.2 Adult acute care beds in St. John's

reduction in the number of acute care hospital beds, and an increase in the number of long-term care beds (1104 in 1981, 1276 in 1993). The number of adult acute care beds in the city is depicted in Figure 1.2.* In 1957, there were ~ 750 beds, and in 1965, ~ 900; this number remained fairly constant until 1990 when there were 941 beds. After this, some of the commission's recommendations became more apparent as bed closures, which at first were implemented only in the summer and at Christmas, became permanent. By 1995, there were 742 acute adult beds (a reduction of 21% in 5 years, while the population of St. John's had increased by 3%). To express the bed numbers in terms of the population of the St. John's region ignores two important factors:

* information courtesy of Eleanor Gardiner and Janet Reid, Newfoundland Department of Health

firstly, ~ one-fifth of the population is made up of children (21% in 1996, markedly reduced from over one-quarter (28%) in 1991) who are not served by adult beds; and secondly, the city hospitals serve people from all across the province. Leaving these considerations aside, the increase in the number of people served by one adult acute care bed from ~ 109 in 1965 to 241 in 1995 (or expressed another way, the reduction from 9.15 to 4.16 beds per 1,000 population) is a remarkable achievement.

In 1961, the provincial length of stay in hospital (LOS) was 11.1 days[4]; by 1990, this had fallen to 8.5 days. With increasing use of day surgery, and new surgical techniques, LOS continued to fall, reaching 5.4 days in 1995. The number of separations rose from ~ 75,600 in 1990, to ~ 78,900 in 1993 (~ 4%), while patient days over the same period decreased by 9.3%. Out-patient and emergency visits have increased but inevitably there is a greater reliance on community services. Beside Community Health, other services which help the elderly and infirm to remain in their own homes include Meals on Wheels, run by VON, and a number of day programs which offer activities, supervision and sometimes therapy for people in need as well as respite for caring families. In principle, shifting more of the responsibility to the community sounds fine; in practice, this approach has not always worked as well as it should, partly because funding to community agencies has not matched the increased workload. The burden on patients and families is sometimes intolerable, such that they can no longer cope, and a new category of hospital admission now exists: community emergency.

This brief summary of health care is concerned only with services in St. John's, and therefore does not describe the magnificent work carried out by doctors and nurses in the outports.

Chapter 2:

2.1 Quick Response Programs have been developed by community health agencies in many regions in Canada to meet the short-term needs of older individuals (and their families) whose ability to function in the community has been jeopardized by illness or injury.

The underlying philosophy is that hospital admission may not necessarily be the most appropriate answer for many of the patients who come to an Emergency Room (ER). Older people as a group are often heavy users of health care resources and as "baby boomers" age, the proportion of seniors in Canada and many other countries is increasing. Hospitalization of the frail elderly often reduces their independence in activities of daily life, and can expose them to infection, as well as causing or increasing mental confusion.[5,6,7] This is not only hazardous for the patient, it is also costly. It is imperative that we consider how to meet the challenge of providing care and support when needed to the elderly and to younger people with chronic problems, while not imposing a crippling burden on younger, more able members of society.

Some people who come to the ER clearly must be admitted; others have relatively minor or non-urgent conditions, which can be dealt with and/or referred to family physicians. Another group of people present with problems which are not so acute as to require admission, yet pose immediate difficulties for the patient and family, such that they cannot sustain their usual routine. Some of these problems can be solved or overcome in a few days or weeks, and given appropriate treatment and support, most patients will return to their previous level of health and independence. Other people have more complex medical or social problems; their ability to function independently may be fragile at best.

Incidents which provoke ER visits by such individuals may raise difficult questions about their future living arrangements. The ER is not conducive to the in-depth discussion needed to arrive at acceptable long-term solutions.

Most Quick Response (QR) programs give care to each patient for a short time only, ranging from 5 to 14 days (or less if appropriate). Some programs accept adults of all ages, but the emphasis is primarily on providing for the elderly. Services (home-making, nursing and other professional care) can be provided within a few hours of a patient leaving the ER, often at an enhanced level by comparison with services routinely available. The aim is to provide a level of support which allows a sick or injured person to remain in familiar surroundings while recovering to a point at which he or she can manage independently (or with no more assistance than previously needed). As soon as the patient no longer requires augmented services, he or she may be transferred from QR to routine community care, or to family care if available and needed. For those patients with multiple problems, or whose health is expected to deteriorate, provision of QR support allows time to assess the situation more fully, and to advise the patient (and family) concerning prognosis and options for ongoing care. Decisions made at home, with a more complete picture of a patient's situation than is sometimes available in the ER, are more likely to meet the needs and acceptance of patient and family.

One of the earliest QR Programs in Canada was set up in Victoria, British Columbia in 1987[8]. Its objective was to provide community-based healthcare quickly to the frail elderly and to adults of all ages with chronic or palliative needs, to maintain their wellbeing and safety at home while avoiding unnecessary admissions to acute care beds.

A QR team consisting of nurses and social workers assessed potential clients in two ERs from 2 p.m. until 10 p.m. seven days a week. Of the 150 patients accepted into the six-month QR Pilot Project, 49% had non-medical "coping" problems; another 56 people met the criteria for QR but lived outside the area or had "unusual living circumstances precluding acceptance by the program" and were therefore admitted to hospital. They made up the control group when the program was evaluated. The authors reported that the ictoria project showed considerable potential for reducing the number of days spent by this group of patients in acute care beds, by comparison with the control group. The service described in the pilot project has since been expanded to include co-ordinating early discharges from hospital as well as receiving referrals from ER.

A trial of QR in Calgary in 1988 emphasized the role of social workers in preventing non-medical admissions[9]. 455 patients over 4 months (11.6% of all those coming to the ER) were referred to the QR project; 95% of these had social rather than medical problems. The target group was adults aged over 70 years, but only one-third of clients referred were over 64 years of age. Only 14 patients required service from home care nurses. The program averted 24 non-medical admissions, and was positively evaluated by essentially all concerned. This report deals with a specific patient group: the role of social work in responding to problems encountered by the wider population of frail elderly or handicapped was not discussed.

In 1991, the Greater Niagara Hospital in Ontario carried out a QR pilot project[10], with the aim of preventing hospital admissions of the frail elderly and disabled adults who came to the ER because their social support networks had collapsed. Of the 237 patients

referred to the QR program during the 12 month project, 19% had social rather than medical problems. The author reported that admission was averted for 206 people who were sent home with appropriate nursing and supportive services. Prior to the introduction of QR, many of these would have been admitted for an average stay of 10 - 12 days. On this basis, it was estimated that over 2,500 hospital days were avoided.

A QR program implemented in Windsor in 1992[11] was evaluated after one year of operation[12]. The program offers immediate nursing assessment, education, counselling and intensive support in the community for up to 5 days to people aged 60 and over who come to the ER. Over the year, 716 patients were referred to QR, of whom 65% had medical problems, and a further 29% had come to the ER because of a fall. Only 6% were judged to have non-medical problems. Many of the people referred had difficulty with activities of daily living (ADL). Twelve percent (89 people) were admitted to hospital. Eighty percent (573/716) were accepted by the QR program and 70% of these received home-making services. Outcome data at 5 days were documented for 528 of 573 patients receiving QR care; follow-up telephone interviews were conducted at 30 days. Only 8% of those accepted by the QR program had to be admitted to hospital. There was no discussion of costs incurred or avoided. Client surveys indicated a high level of satisfaction with services provided.

The QR projects described so far were evaluated positively, in some cases enthusiastically, by care recipients and staff members. In each case, it seems that the program was designed and implemented first, with evaluation occurring as an afterthought. With the exception of the Windsor-Essex program, evaluation was carried out by staff

involved in program delivery and claims of potential savings were based on dubious assumptions. In-house evaluation of service delivery is frequently carried out as part of a reporting mechanism to administrative or funding bodies and therefore it is in the interest of those assessing the program to highlight the achievements. The object of the exercise is to justify continuance of a program and if possible, an increase in funding. This process differs fundamentally from true evaluation, in which underlying principles are that clear questions should be formulated and that limitations and potential biases should be recognized and as far as possible, minimized or avoided.

Ideally, in evaluating the need for a new service, or efficacy of an established program, careful analysis and valid comparison are made, *e.g.* management and costs for patients who fulfil target criteria before and after implementation of the service. Unfortunately, a problem with this approach is the current speed of change in health care delivery. It is difficult to compare costs and outcomes when there are factors the investigator cannot control, and it is not always easy to find all the data needed if part of the study is retrospective. A better method of comparison is to use random assignment of patients to a control or intervention group, with standardized collection of data.

In only one of the studies quoted here was a control group identified (Victoria), but it was a "naturally occurring" group, flawed by selection bias, differing in area of residence, and in some cases, in social circumstances. Subjects in the control and intervention groups were well matched for age, but no information was provided concerning diagnosis or severity of disease, so it was not possible to judge whether the two groups were comparable.

In the absence of a properly selected control group, it is difficult to be certain that avoidance of hospital admission is a consequence of the QR intervention. In all the reports quoted, unfounded assumptions and claims were made regarding the number of admissions averted: in Victoria, 150 admissions over six months; in Calgary, 24 over four months; in Niagara, 206 over one year; and in Windsor-Essex, more than 500 over the course of one year. The underlying assumption in each of these reports is that all of the patients referred to QR would in the absence of the program have been admitted. Assumptions were also made concerning the length of stay (LOS) which would have been associated with these admissions: in Victoria (by using the mean LOS of the non-comparable control group) 12.5 days; in Calgary, 8.4 days (based on mean hospital LOS); in Niagara, 12 days (based on previous experience); in Windsor-Essex, 8 - 9 days (573 patients admitted to the QRP, and the stated saving of 5,000 bed days). While some of these estimates may have been reasonable best guesses, the claims made for saving of patient days in hospital are not justified by the evidence presented.

It is not clear that data were collected consistently in these studies: the authors of the Victoria study reported that data were "descriptive and specific to context". In the other studies, information was documented as required by the service agencies. In only two studies (Niagara and Windsor-Essex) was there any mention of patients returning to the ER with the same or similar complaints within a short time of enrolment.

Apart from a brief comment in the Calgary report mentioning that having a service provider (the social worker) conduct the follow-up survey might prejudice the opinions

expressed about the quality of service provided, none of the reports discussed limitations of their studies. The Windsor-Essex evaluation noted that some clients might benefit from more than the 5 days of QR intervention offered, and suggested that research be continued to clarify this issue.

Evaluation of these QR programs could have been improved had criteria been set up at the outset, together with the measurement tools to be used. Random selection of intervention and non-intervention groups, treated identically with respect to data collection provides the best basis for judging efficacy of intervention.

In 1992, a feasibility study was carried out in St. John's Newfoundland to see if numbers of appropriate patients were sufficient to make a QR study in St. John's practicable (see page 20). The Quick Response Project was planned in 1994 as a research-oriented initiative to develop, operate and evaluate a QR Team. As described in the following chapter, the proposal was for a randomized clinical trial designed to overcome some of the problems noted above in assessing the efficacy of QR Programs.

Subsequently, we learned of three other QR initiatives: a study conducted in 1991 assessing the role Home Care might play in reducing hospital admissions through the ER, conducted by staff of the Wellington-Dufferin-Guelph Home Care Program[13]; the Kingston, Frontenac, Lennox and Addington Quick Response Service (QRS)[14] which was implemented in September 1994; and the Ottawa-Carleton QRS[15] implemented in October 1994.

The Guelph study was in two parts: a retrospective analysis of 50 charts, randomly selected but in proportion to annual admissions to each hospital service, of patients who had been admitted through the ER and discharged from hospital in October 1990, and a prospective review over 5 days (Monday to Friday) in March 1991 of all 63 admissions through the ER. The authors concluded that 10/113 (9%) of the total sample of hospital admissions could have been avoided by appropriate use of existing Home Care services. The other 91% of patients required admission for monitoring or treatment, and it was felt that even if a QR program had been available, these admissions would not have been averted. This was a carefully conducted study with clear research questions posed and assumptions stated. Limitations noted by the authors include: no control group; very limited time frame (avoiding the weekend in the second part of the study); inconsistent information. They made several recommendations for further study, but concluded that with greater awareness and more effective use of the existing Home Care Program, it was not necessary to develop a Quick Response Program.

The Kingston program ran for one year, providing up to 5 days intensive home support for 123 patients. Over 9 months, an estimated 41 admissions (out of 71 referred) were averted, although without a control group it is difficult to substantiate this claim. The study suffered from inadequate enrolment. The financial implications of the program received careful analysis, and while all parties agreed that the QRS was providing a necessary service, it was concluded that it was not economically viable in a time of financial restraint.

A brief summary of the Ottawa QRS shows that in one year, 195 patients were referred and 121 accepted. There was no control group. Using Case Mix Group average LOS data, the authors concluded that had their subjects been admitted to hospital, they would have had an average LOS of 6.79 days. Total annual QRS operating costs were $233,882, and averted hospital costs were $383,690, for a per patient savings of $1120. The logic for the calculation appears sound, but the presumption that all 121 patients would have been admitted to hospital in the absence of QRS is not.

A brief outline of the QR programs described here is presented in Table 2.1.

In summary, consideration of the need for a program, and evaluation of a program already in place are necessary components of health care planning. Criteria must be established, preferably beforehand, so that need for a program, or its operation can be judged without bias. An annual report detailing statistics and achievements of a health care unit is not a substitute for an objective review. Clear questions, carefully selected and worded, provide the basis for objective evaluation.

Table 2.1. Summary of some Quick Response Programs in Canada

Program, date	Duration (months)	Clients accepted (n)	Health status reported	Evaluation		
				Averted unnecessary hospital admissions	Control group	Cost analysis provided
Victoria BC[8] 1987	6	150 + 56 controls	examples of presenting diagnoses, no outcomes	yes, assumed to be 150	yes, but neither randomly allocated nor comparable	partial
Calgary AB[9] 1988	5	455	no	yes, 24 non-medical	no	no
Gr. Niagara ON[10] 1991	12	237	no	yes, estimated 206	no	partial
Windsor ON[11,12] 1992	12	573	brief summary given	yes, not stated how many	no	no
KFLA ON[14] 1994	9	123	precise listing of primary diagnosis	yes, estimated 41	no	yes
Ottawa ON[15]	12	121	summary by diagnostic group	assumed to be 121	no	yes

Chapter 3

3.1 Rationale: recognizing the need to increase the options for caring for the elderly in St. John's, we decided to investigate further the Quick Response concept developed in other centres. Adults who present with problems which are neither so severe as to require immediate admission to hospital, nor so minor that further treatment is unnecessary or not urgent, may appropriately be managed by a Quick Response Program. As described in the previous chapter, several Quick Response Programs have been put in place in Canada, on the assumption that the number of admissions to hospital could be reduced, and health outcomes would be at least as good, while the cost of providing treatment would be substantially reduced. It was not possible from the evidence provided to reach firm conclusions concerning the efficacy or financial implications of these programs and we decided to conduct a randomized clinical trial of QR, involving Community Health - St. John's Region and the three adult general hospital Emergency Departments in the city.

3.2 Feasibility Study: a feasibility study was carried out in 1992, by looking at the spectrum of patients aged 65 years and over who attended Emergency Departments at the three adult general hospitals, over a period of one month each: May, June and July respectively. Of the 8412 visits made to adult ERs, 817 were made by 677 patients aged 65 years and over, who were resident in the St. John's area. An expert panel of physicians and nurses concluded that of these patients, 75 (11%) might have benefited from QR had it been available. Thirty-three (44%) of these patients were admitted to hospital at the time of their first ER visit, and another 12 (16%) were admitted

subsequently for management of the initial complaint. The group made 52 repeat visits to ER in the 3 months of follow-up; only 3 of these repeat visits were for new reasons. On the basis of these results, it was concluded that there were sufficient patients attending the city ERs who met age and residency requirements to make a trial of QR feasible.

3.3 Study Proposal: a proposal to develop, operate and evaluate a Quick Response Team was submitted to Health Canada - National Health Research Development Program late in 1993, and after clarification of some details, was approved in the summer of 1994. The primary hypothesis of the study was as follows: QR will cost less than standard care while proving as effective as standard care with respect to (a) patient health status, (b) caregiver burden, and (c) patient and caregiver satisfaction with services provided. The sample size for the trial was estimated on the basis that QR would only be viable if it was associated with a reduction in the cost of care funded by the provincial government. To be sure that the extra costs associated with running QR could be funded without increasing the overall health budget, it was felt necessary to demonstrate that inpatient days could be reduced by at least 25%. The proposal was for a study large enough to detect a QR-associated 25% reduction in the average length of inpatient stay during a three-month period following initial presentation, with a power of 80% and a one-tailed type 1 error rate of 5%. This would require a total sample size of 630 patients divided into two randomly allocated equal size groups, one group to receive QR intervention and the other to receive standard care. The anticipated length of the study was 8½ months, plus a pilot phase of 6 weeks.

3.4 Planning: medical and nursing representatives from the three adult Emergency Departments, and nursing and administrative representatives from Community Health met with the researchers to plan the practical details of running the trial. Eligibility criteria[1] for study subjects included: presentation to the ER at the General Hospital Health Sciences Complex, the Salvation Army Grace General Hospital or St. Clare's Mercy Hospital; residence in the area covered by Community Health - St. John's Region; willingness to take part, to answer questions and to allow the researchers access to medical records; a problem or constellation of problems which, though demanding attention, would not absolutely require hospital admission if care could be provided in the community more rapidly and at a greater level than routinely available under existing Community Health programs. The acute reason for presentation to the ER had to have the potential for resolution within the two weeks allowed for QR intervention, even though chronic underlying problems might still remain, *e.g.* a urinary infection in a patient with dementia, or a fracture in a patient with terminal disease.

Orientation sessions were given to Emergency Room staff by the principal researchers (BB and DN) before the QR project commenced, and ongoing contact was provided by the QR assessment nurses who visited the ER departments once or twice daily to advise *re* suitability of patients for enrolment in the trial.

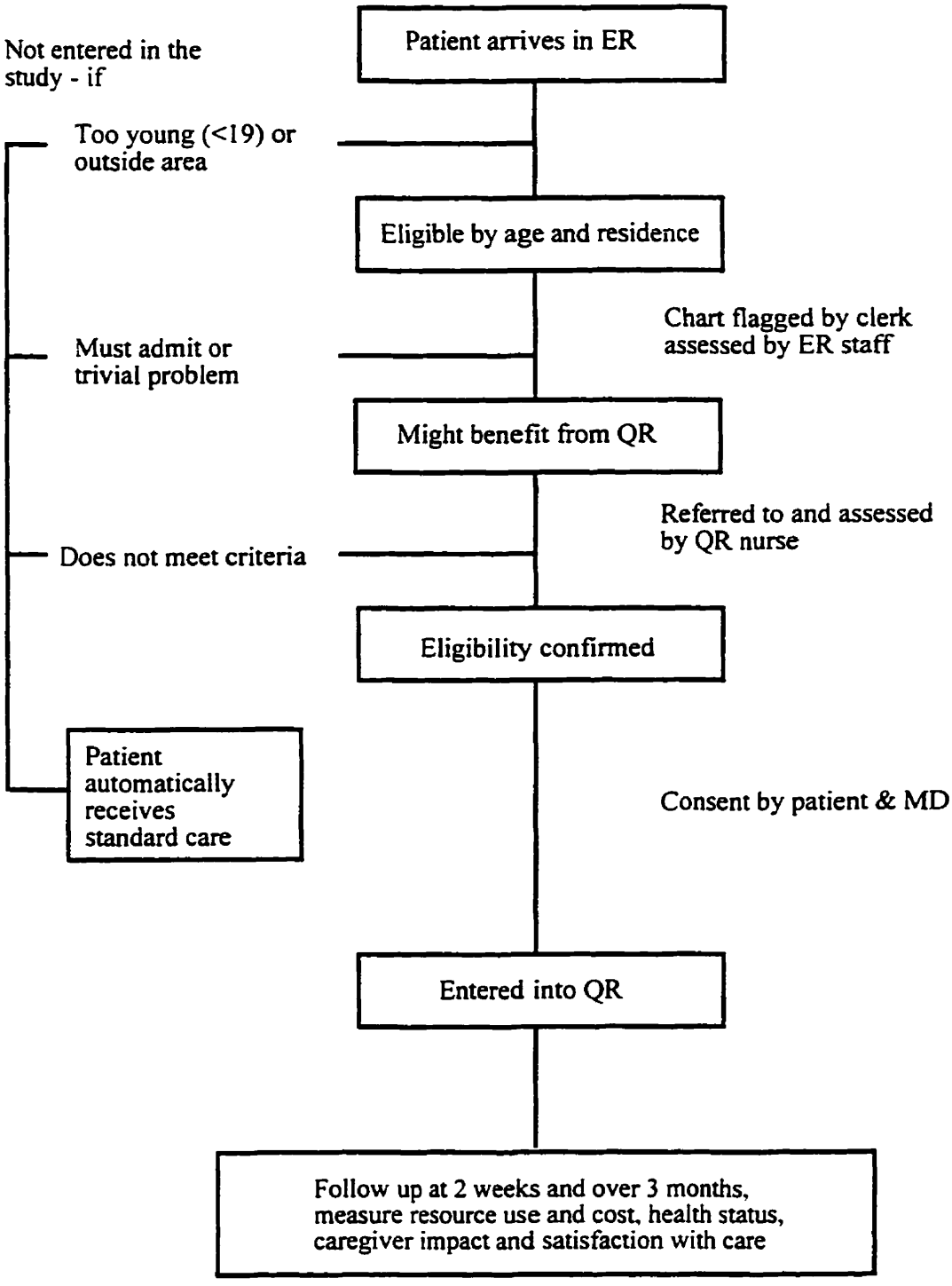
---

[1] Criteria for admission to the study are listed in the appendix, page 127.

3.5 Pilot Phase: the pilot phase of the QR Project, initiated 1995.02.17, was used to identify operational problems, and to develop the Satisfaction Questionnaire. Initially, only patients aged 60 years and over were deemed eligible for entry to the study, but during the pilot phase it was decided to offer participation to all patients aged 19 years and over who met study criteria. Once the ER physician, or a consultant who had seen the patient at the request of the ER physician, identified a potential subject, the QR assessment nurse was notified, and the process of confirming eligibility was followed as outlined in the flow chart, Figure 3.1. The physician decided whether the patient would in the absence of the QR project be managed by admission to hospital, or by treatment in the ER (if required) and/or follow-up in the community. The purpose and format of the QR study were then explained to the patient (and if appropriate the caregiver or accompanying person) by the QR nurse, and written information describing the study was provided before they were asked if they would consent to take part. The choices available to them were explained as clearly as possible by the QR nurse. If they agreed, the patient (and if appropriate, the caregiver) signed the consent form(s). The usual demographic details were recorded on the Community Health QR registration form, together with the physician's diagnosis and orders and the QR nurse's analysis of patient problems. In the pilot phase, all eligible patients who consented to enter the study received QR intervention.

Figure 3.1. Protocol for QR study Pilot Phase

For approximately 6 weeks, starting on 13 February, 1995



Those patients who declined to participate in the study received ongoing care as ordered by the attending emergency physician, and arrangements were made by ER staff in the usual manner.

3.6 Randomization (Evaluation) Phase: in the evaluation phase, patients who consented to enter the study were randomized to receive either standard care (admission to hospital or care in the community, as decided by the physician) or QR. Randomization was carried out using computer-generated random numbers, in blocks of six. The process of determining eligibility, and the subsequent randomization into Standard Care (SC) or Quick Response (QR) groups is shown in Figure 3.2.

3.7 Aims of QR: by putting appropriate community services in place as soon as need is identified,

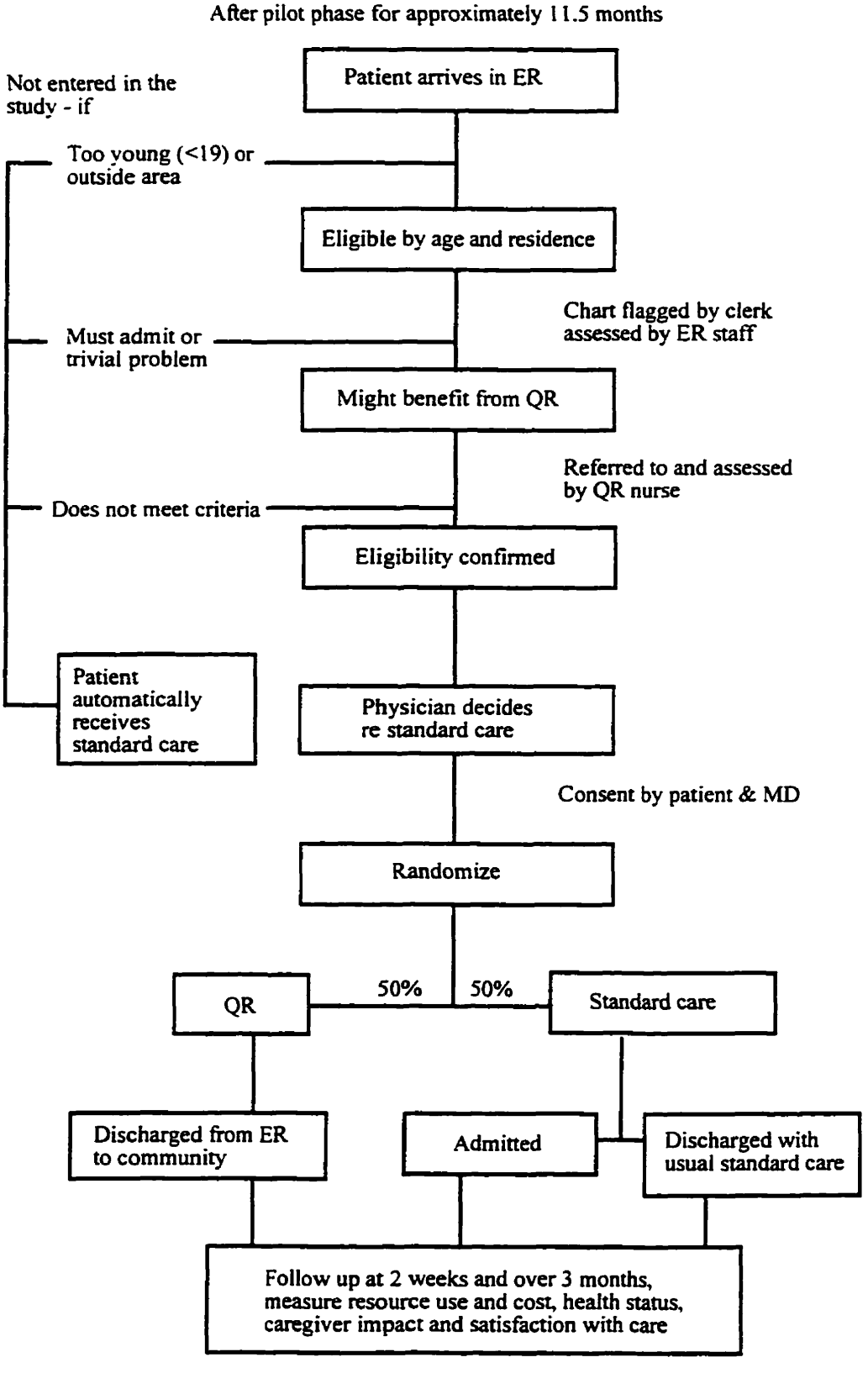
(a) unnecessary hospital admission may be avoided

(b) outcomes will be at least as good and in many cases better and/or achieved more quickly

(c) the cost of providing health care will be less than providing the care routinely available at present

(d) caregiver stress will be no greater than is experienced under the current system and in many cases will be less.

Figure 3.2. Protocol for QR study Randomization Phase



All subjects who entered the trial completed a number of questionnaires (described on pages 28 - 35) designed to measure degree of independence, ability to cope with everyday activities, and the impact of illness on various aspects of daily life including socializing and mental health. These questionnaires were administered on three occasions: firstly, in the ER or as soon as convenient thereafter, then two weeks after the ER visit, and finally three months after the visit.

If there was a caregiver, he or she completed two questionnaires designed to measure some of the stresses inherent in the caregiving role. When the 2-week visit was made, an additional questionnaire was presented, addressing patient and caregiver satisfaction with the care provided in the 2 weeks after enrolment.

The care received by subjects in the two weeks and three months following enrolment in the trial was documented as fully as possible, from information provided by the subjects and their caregivers, from reports by Community Health staff, and from records of physician services provided by the Newfoundland Medical Care Program (MCP) . An estimate was made of the cost of services provided to each patient in the study. Patients or their families were asked to fill in a log to record any care provided or visits to hospital or to a doctor's office: this helped to provide confirmation of services provided. Specific permission to access the medical records and MCP data was requested and granted in the consent forms signed by study subjects.

3.8 Questionnaires used in the Quick Response Trial: no single questionnaire can be expected to capture the complex nature of human reactions to illness or accident, or the wide range of ability encountered when working with the diverse population of patients who attend ERs. For this reason, three well-tested questionnaires (the modified Barthel Index, hereafter referred to as the Barthel), the Sickness Impact Profile (SIP) and the Short Form-36 (SF-36)) were chosen to reflect aspects of function, health status, and the impact of sickness on an individual's life; two (the Relatives' Stress Scale (RSS) and the General Health Questionnaire (GHQ-30)) were chosen to measure some of the objective and subjective aspects of caregiving. Satisfaction with care provided is a difficult concept to measure, and no instrument directly addressing satisfaction with health care both at home and in the hospital could be found in the literature. One of the investigators (DN) developed a short questionnaire for patients, and where appropriate a version for their caregivers, to measure satisfaction with health care and home-making support.

Many scales used in health care research, including several of those listed above, are ordinal: a scale is constructed by selecting items which relate to a parameter of interest, and assigning a rank order to the possible answers. From this a numerical score is obtained. The authors of the Barthel and SIP have assigned weights to the various items and responses, based on analysis of their results; the General Health and Pain scales of the SF-36 also have some weighted items; if one assumes that the populations they studied are comparable with the population to be tested, this weighting can be accepted. However, other scales, including the RSS and most of the SF-36 are clearly ordinal. Two important concerns must be recognized. One is that the conceptual distance between different items on the scale,

and between possible reponses, may not be known and therefore the sum totals obtained cannot be (but sometimes are) subjected to statistical methods appropriate to ratio and interval scales[16]. Another important concern, common to all scales, is that it is difficult to define what difference in numerical score constitutes a clinically important difference: this is probably not consistent over the range of possible scores: a numerical gain of 4 points at one end of the scale may indicate a greater improvement in function than a gain of 5 points in the middle of the scale[17]. Thus the meaning of a particular gain depends to a considerable extent on the baseline score. Items in the middle of a scale, *e.g.* the Physical Function of the SF-36 scale, are often closer together than at either extreme of the range, so it may be more difficult to demonstrate change in subjects who in terms of the parameter to be measured lie at the upper or lower limits of the scale. Rasch analysis[18,19] provides a way of dealing with ordinal data, by estimating the conceptual distance between items on a hierarchical scale, using the mean of the natural log of the odds of the "average" patient making the transition from one category to the next higher one on the scale. This gives a value expressed in log-odd units, "logits", which define the location of that category on the scale, so that the distance between items is known for the test population. This would be appropriate for the Physical Function and Mental Health scales of the SF-36 and for the RSS and GHQ.

Alternatively, it is possible to use scores obtained from large samples of clinically defined patients to obtain a pragmatic understanding of what scores represent in those populations: if the sample described matches the group to whom this information will be applied, this will give a reasonable idea of a clinically significant difference.

For each of the scales used in this study, literature reports have been used to identify scores tied to clear clinical differences; we shall use these differences to represent significant clinical change.

3.8.1 Modified Barthel Index[20]. This instrument measures independence in 14 basic items of daily living (ADL) and takes approximately 5 minutes to answer. It does not address social interaction, or mental health except insofar as it impacts on physical function. Independence in self-care is scored out of 53, and in mobility out of 47, for a total of 100% indicating complete independence; zero represents absolute dependence. The Barthel Index[21] was shown to be valid, reliable and sensitive when used for evaluation of severely handicapped individuals[22].

The original index was modified by distinguishing two levels of independence: intact-independent and independent with difficulty, or with an aid such as a cane. The modified version has been shown to be a valid measure of physical impairment, with the ability to discriminate between differing levels of need for assistance in the tasks of daily living. Granger[22] used arbitrary cut-off points of 20%, 40% and 60% to represent different levels of ability in a population of severely handicapped adults. The scoring system indicates that someone who had difficulty in all areas, yet could manage independently, would score 94%, while someone who required help with transfers and toileting would score 87%: this suggests that 7% represents a real difference in function.

3.8.2 Sickness Impact Profile (SIP)[23]. This questionnaire consists of 136 statements which address the impact of sickness on many aspects of daily life such as physical function, socializing and family relationships, mental health and degree of orientation. Only those statements which accurately describe the subject on that day, and which are the result of illness, are checked. If no statements are checked, the score is 0%, indicating that an individual perceived no impact by sickness on his or her daily life. There are twelve SIP subscales, two dimension scales (physical and psychosocial) and an overall total score; only the physical dimension and total scores will be discussed here. Comparison with the American Rheumatology Association classification system shows that people with rheumatoid arthritis with complete ability to carry on all usual activities without handicap (Class I) scored 8% overall (physical dimension 6%); those with pain and/or limited joint range but adequate function (Class II) scored 15% (physical dimension 13%); those whose disease prevented them from pursuing their regular occupation or daily activities (Class III) scored 20% (physical dimension 19%), and those who were largely or wholly incapacitated (Class IV) scored 26% (37%)[24]. Another study of 87 individuals with acute back pain which improved over the three weeks between two dates of examination reported a mean change in SIP total score of 7.5% (physical dimension, 10%). The SIP has been shown to be valid[20] and reliable, but some concern has been expressed in the literature that the SIP may be less responsive to improvement than to deterioration[25],[26]. In a study of 54 patients undergoing hip arthroplasty, brief health status measures such as the SF-36 (physical subscale) were shown to be equally or more sensitive to change.[27] A disadvantage of the SIP, particularly when used with a sick or elderly population, is its length. It takes ~ ¾ hour to complete, sometimes longer, and many respondents find it very tiring.

3.8.3 Short-Form 36 (SF-36)[28]. This is a much shorter instrument consisting of 36 questions relating to physical function, pain and its effect on activities, depression and anxiety, socializing, and perception of health. It takes ~ 10 - 15 minutes to complete. Perfect function would give a score of 100% in each of 8 areas: physical functioning (PF); role functioning - physical (RP); bodily pain; general health (GH); vitality; social functioning; role functioning - emotional; and mental health (MH). The SF-36 has been shown to be reliable across diverse medical and psychiatric groups with Cronbach's α for PF being 0.93 and for MH 0.90[29]. The questionnaire has content, construct[30] and discriminant validity[31], and is reliable when used in a general population[29]. When used to follow workers with musculoskeletal disorders, the PF and pain scales were found to have adequate reliability and to be more responsive to improvement than the SIP physical dimension or total score. The pain, RP and SF scales are predictive of hospitalization within the following 2 years, and the GH and PF scales have predictive validity in terms of mortality over the following 4 years. Based on a study of over 9,000 adults (mean age 46 years) visiting family physicians in 3 U.S. cities[32], the mean PF score of someone with no chronic conditions was 86.0; of those with "arthritis" or back problems, 77%; patients with chronic respiratory problems scored ~ 73%, and patients in congestive heart failure ~ 63%. In another study, the mean PF score of 185 healthy older adults (mean age 72.5 years) was 77.1 ± 22.4[33]. Patients with minor medical problems scored 29.7% more on the MH scale than patients with psychiatric conditions[30]. SF-36 General Health Perception Scale group means are reported in the Beaver Dam Health Outcomes Study:[34] differences of ~ 7.5% distinguish between groups of people with and without colitis

or orthopedic problems, and 10% between those with and without angina.

Floor and ceiling effects have been noted with the SF-36 as with other measures of health: in a sample of chronically sick adults, aged 75 years or more, minor floor effects were seen in the PF scale (3.1%) and none in the MH scale. Ceiling effects were seen in 5.6% of respondents in the PF scale, and 11.1% in the MH scale. Much larger effects at both extremes were seen in the role-physical and role-emotional scales, and at the ceiling in the pain and social function scales. Skewing of the score distributions makes these latter scales more problematic when used across age groups or in the elderly[35].

3.8.4 Relatives' Stress Scale (RSS)[36]. This scale was designed by Greene and colleagues to measure the stress experienced by caregivers looking after a demented elderly relative at home; it has also been used for caregivers of stroke victims[37]. It consists of 15 questions about many of the stresses well known to caregivers, and takes 5 - 10 minutes to complete. Based on data from a small sample of caregivers of demented elderly people, the authors proposed three subscales: personal distress, life upset and negative feelings. The subscales are added to give a total score ranging from 0 - 60; scores below 10 reflect mild or minimal stress; scores above 30 indicate considerable distress. The authors reported test-retest reliability of 0.85; construct validity for the scale was supported by concurrent use of their Behaviour and Mood Disturbance Scale, reported in the same paper, and for the dependent family member, measures of cognitive function,[38] ADL and self-care[39]. They did not propose a link between the objective demands placed on caregivers and their subjective reactions.

Eagles *et al.*[40] used a modified scoring system for the RSS (a 3-point scale

rather than the 5-point scale used by the authors) in a study of co-resident supporters of 79 elderly subjects living in the community. They demonstrated a significant difference ($p = 0.001$) in the degree of stress experienced by supporters of non-demented people compared with supporters of moderately or severely demented people; the mean difference in group scores was 6.1 points. Scores ranged from 0 - 12 for caregivers of the non-demented, (group median score 1.71, mean 2.7) and 0 - 20 for caregivers of the demented (median for caregivers of the demented 5.75, with mean scores 5.5 for caregivers of people with mild dementia, 7.5 and 10.6 for caregivers of those with moderate and severe dementia. Even though they were clearly under more stress than those caring for non-demented folk, supporters of demented people did not show any greater degree of psychiatric distress (as measured by the GHQ-60, see below).

3.8.5 General Health Questionnaire (GHQ30)[41]. This instrument belongs to a series of GHQ instruments of varying length and was chosen to give an indication of psychological well-being in caregivers. It has been widely tested in many different populations and has been shown to be valid and reliable in measuring psychological distress, while not being excessively tedious to complete (usually ~ 10 - 15 minutes). The dichotomous scoring system recommended by Goldberg was designed to screen the general population for minor psychiatric morbidity[42], and gives a cut-off of ≥ 5 as representing significant individual psychiatric distress. Goldberg did not claim that this system could measure the mental health of those identified as cases, although other authors have suggested that this might be possible with alternative scoring systems[43].

3.8.6 Satisfaction Questionnaires. These questionnaires were developed during the pilot phase of the study, and used during the randomization phase. Fourteen questions address aspects of care previously identified as important by a convenience sample of professional caregivers. Areas covered included skill, empathy, willingness to answer questions and provide explanations, flexibility and reliability. Questionnaires were answered by patients, and when applicable, by their (non-professional) caregivers, using the 7-point Delighted - Terrible Scale[44]. There were a few open-ended questions designed to elicit comments about aspects of care which pleased clients or fell short of expectations. The satisfaction questionnaires took ~ 10 minutes to complete.

Chapter 4 - Results

Enrolment of subjects was substantially less than anticipated, at most 9 subjects per week (see Figure 4.1). This was achieved towards the end of the pilot phase, probably in response to one of the investigators (BB) visiting each ER to explain again the rationale for the study. Thereafter, enrolment declined and in June and July was only 2.2 per week.

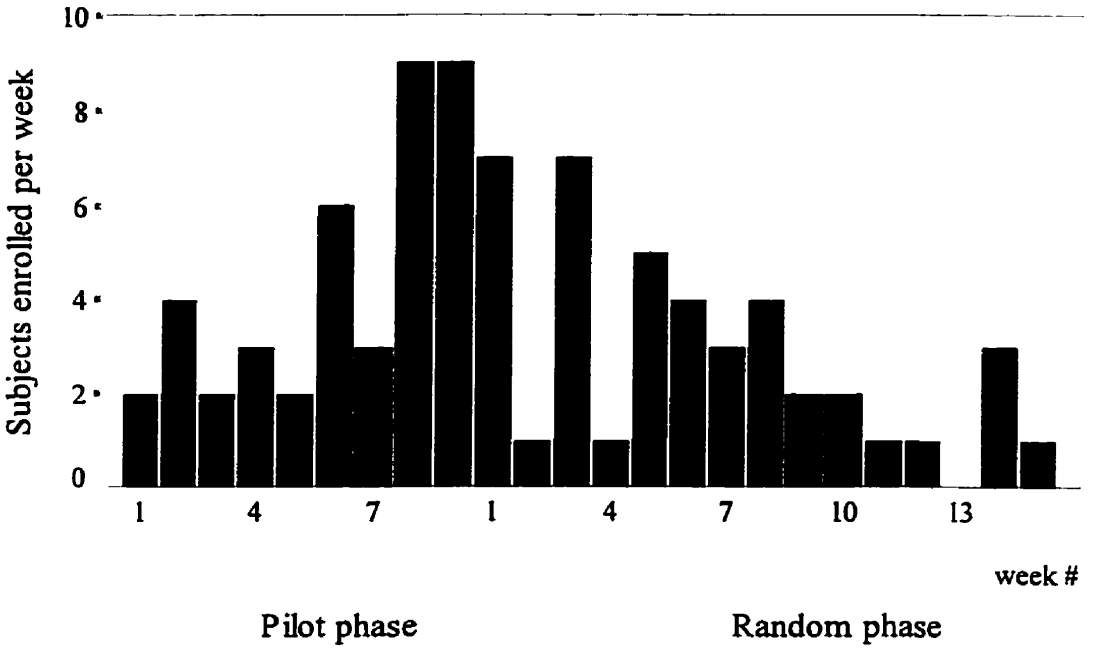


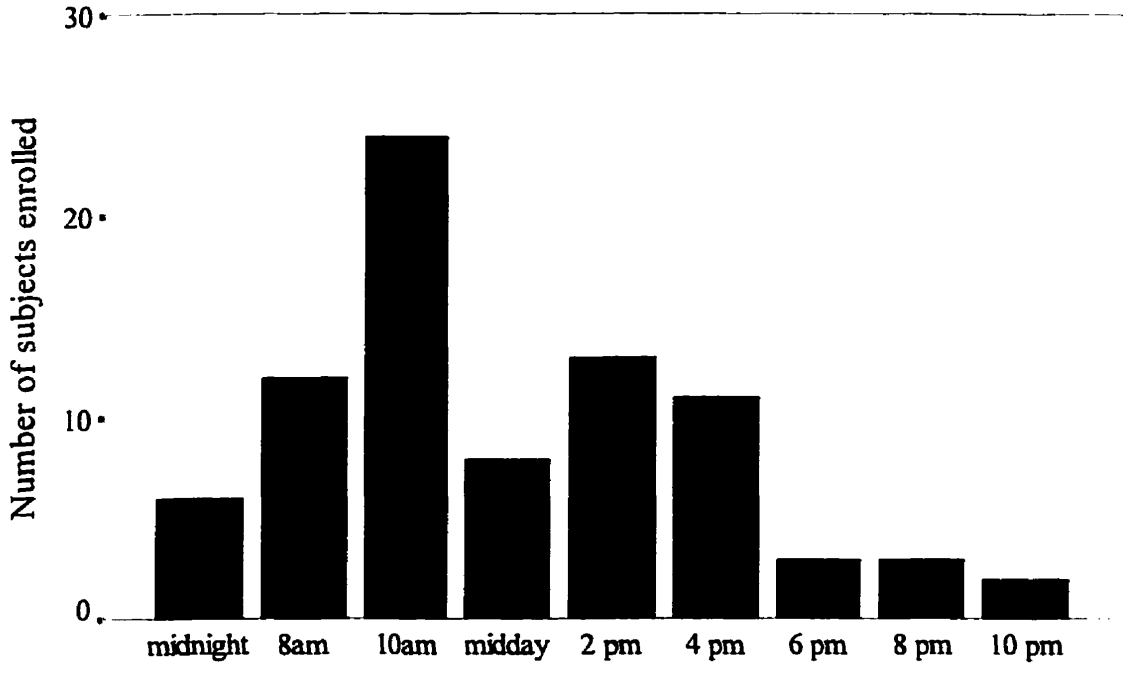
Figure 4.1 Study subjects: enrolment per week

Little or no improvement in numbers was expected over the summer months (because fewer patients are registered at the ER during the summer; staff have less time than usual because many of them take annual leave. In the 24 weeks of the study (pilot and random

phase) an average of 3.6 subjects per week were enrolled. Given this rate of recruitment, it would have taken 3½ years to enrol the projected number of subjects. The cost of continuing for this length of time (~ 4 times the anticipated duration) would have been prohibitive, and therefore the study was terminated after only 24 weeks. Possible explanations for the discrepancy between predicted and actual enrolment are discussed on pages 74 - 77.

The task of tracking study subjects, the care they received and the associated costs, was labour-intensive. Confirmation of information and dates given by subjects and their families had to be sought from a number of different agencies, including hospitals, Community Health and MCP. It was not possible to access information about medications dispensed. In the future, linked data bases will make it easier and quicker to access such information: the Newfoundland and Labrador Centre for Health Information was set up in 1997 to implement the linking of health data to unique personal identifiers, while at the same time protecting patient confidentiality.

The Pilot Phase lasted 9 weeks, 95.02.13 - 95.04.16; 57 potential subjects were referred to the QR project of whom 40 were enrolled. The Evaluation (Randomization) Phase ran for nearly 15 weeks, 95.04.17 - 95.07.28; 57 people were referred and 42 were enrolled, of whom one had already taken part in the Pilot Phase. The presenting problem initially was a fracture, caused by a fall; in the Randomization Phase, her underlying problem had not changed, but she presented with an infection. In analyses using data from both phases of the study, this individual is represented only once unless otherwise stated.



Time blocks shown here depict midnight to 8 a.m., 8a.m to 10 a.m.
10 a.m. to midday ..... 10 p.m. to midnight

Figure 4.2 Emergency Room arrival times

Most of the patients enrolled in the study arrived in the ER during the day: 44% (36/82) came between 8 a.m. and noon, and a further 39% (32/82) between noon and 6 p.m. (see Figure 4.2). The mean time between ER arrival and the page to the QR nurse was nearly 4 hours (3 hours 52 minutes), but this includes 7 people who were kept in the ER overnight before being referred. Excluding these subjects, the mean lapse of time between arrival and the summons to the QR nurse was just less than 3 hours. A QR nurse was on duty from 8 a.m. until 10 p.m., covering the 3 ERs, and mean response time was 41 minutes (median 30 minutes); mean QR assessment time was 61 minutes (median 51).

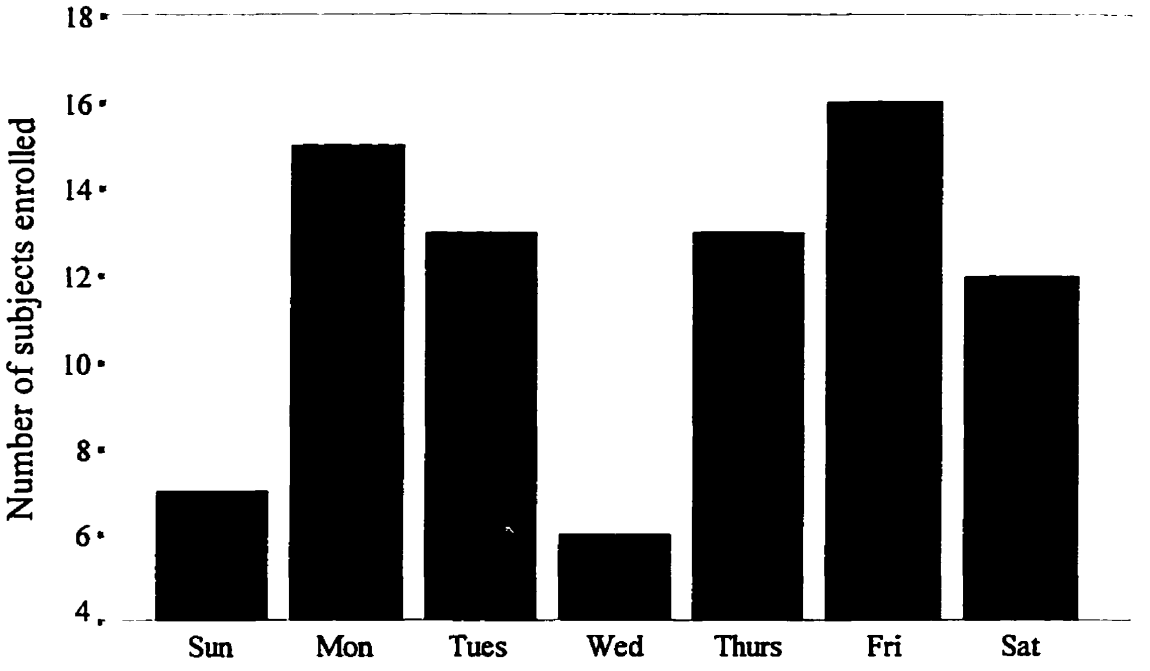


Figure 4.3 Study enrolment by day of the week

Enrolment was disappointingly low, with Wednesdays and Sundays attracting the lowest numbers (Figure 4.3).

One hospital (St. Clare's Mercy Hospital, SCMH) accounted for over 50% of the subjects enrolled in the study (19 people in the Pilot Phase and 25 in the Randomization Phase): this is explained in part by a strong orthopedic department (26/44 subjects at this hospital presented with fractures or soft tissue injuries) and in part by the strong support given by senior nursing staff.

Twenty-nine percent of all subjects (11 people in the Pilot Phase and 13 in the Randomization Phase) came to the Health Sciences Complex (HSC); 17% (10 and 4 people in each phase) attended the Salvation Army Grace General Hospital (SAGGH). With the exception of the 13 patients who entered the Randomization Phase of the study *via* the HSC ER, who were somewhat younger (59 $\pm$ 23, median 60 years), there was little difference in the age of patients enrolled at the three hospitals, ranging from means of 70 to 78 years, medians from 74 - 78 years.

As described in Chapter 3, all subjects who agreed to take part in the Pilot Phase received QR care. Based on the ER physician's decision as to standard care (in the absence of QR, would or would not admit to hospital) they were assigned to Group 2 or Group 4, see Table 4.1. In the Randomization Phase, after the physician had decided what standard care would be, and patients had consented, randomization then assigned "would admit" subjects to Group 1, Standard Care (SC) *i.e.* admit to hospital, or to Group 2, QR at home. Similarly, "would not admit" subjects were randomized to Group 3, SC *i.e.* routine care at home as requested by the ER physician, or to Group 4, QR at home.

The original intent was to compare Group 1 with Group 2 subjects, and Group 3 with 4, but because there were so few subjects in the "would admit" category, randomized SC Group 1 was amalgamated with Group 3, and QR Group 2 with Group 4 for the purpose of analysis.

Table 4.1. Enrolment by Group: Pilot and Randomization Phases

Pilot Phase				
Group	**1**	**2**	**3**	**4**
n		9		31
Sex		1 ♂, 8 ♀		14 ♂, 17 ♀
Age in years mean ± SD (median)		64.4 ± 17.9 (72.0)		73.1 ± 15.3 (76.0)
Random Phase				
n	3	2	17	20
Sex	2 ♂, 1 ♀	1 ♂, 1 ♀	7 ♂, 10 ♀	11 ♂, 9 ♀
Age in years mean ± SD (median)	50.0 ± 27.8 (45.0)	65.0, 74.0 (70.0)	69.8 ± 19.7 (79.0)	68.3 ± 21.1 (77.0)

Group 1: Standard Care, admitted to hospital.
Group 2: QR at home: in the absence of QR would have been admitted to hospital
Group 3: Standard Care at home
Group 4: QR at home, in the absence of QR would not have been admitted to hospital

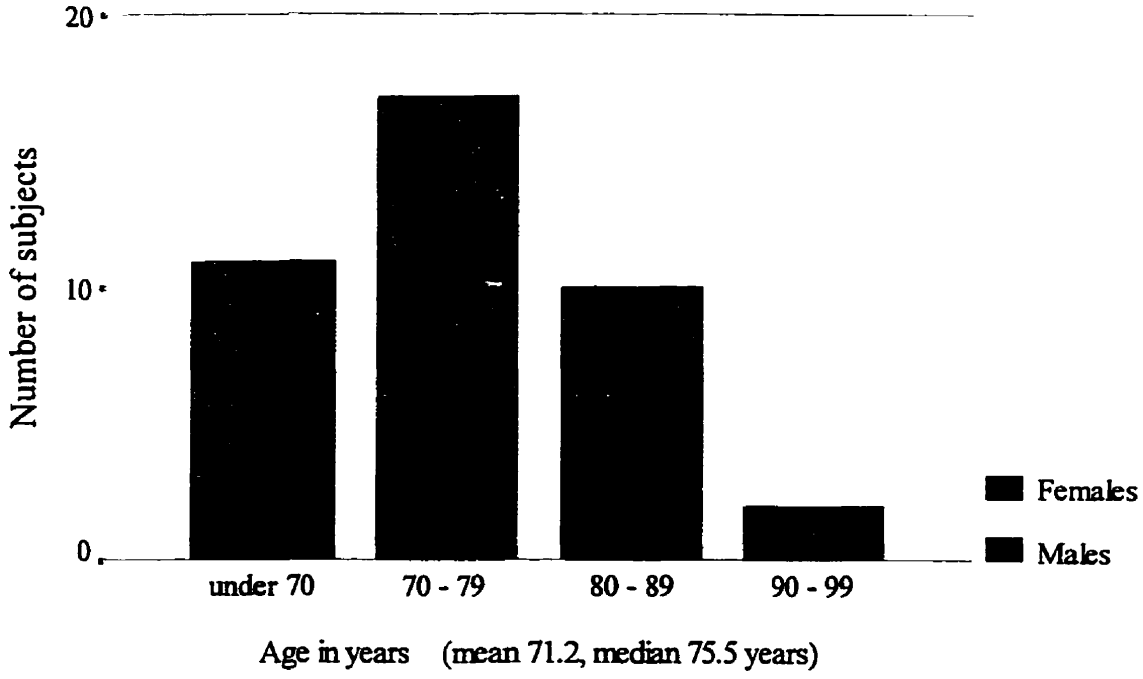


Figure 4.4a Pilot Phase subjects by age group and sex

Figure 4.4a shows the age group and sex distribution of subjects in the Pilot Phase (who all received Quick Response). Figures 4.4b and 4.4c show age and sex distribution in the two randomized groups. Randomization to SC or QR was carried out in blocks of six, after consent was given, and worked reasonably well, given the relatively small numbers involved. The age difference between randomized SC and QR subjects was 1.6 years, not significant either clinically or statistically (Student's 2-tailed t-test $p > 0.25$).

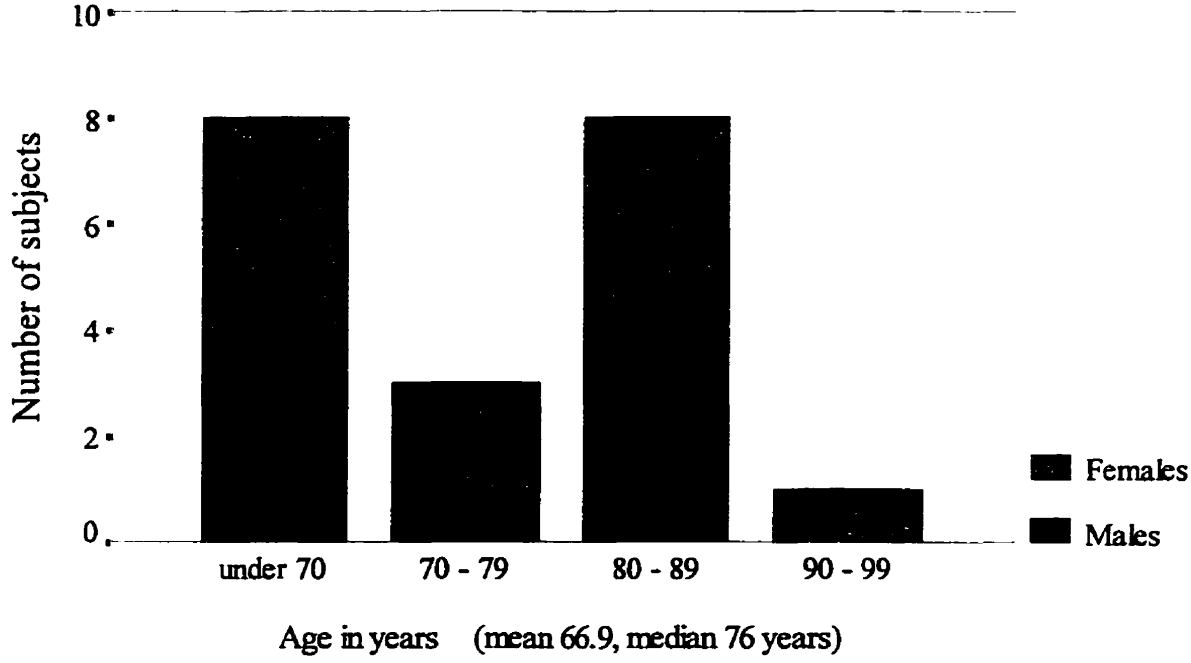


Figure 4.4b Randomization Phase subjects, Standard Care, by age group and sex

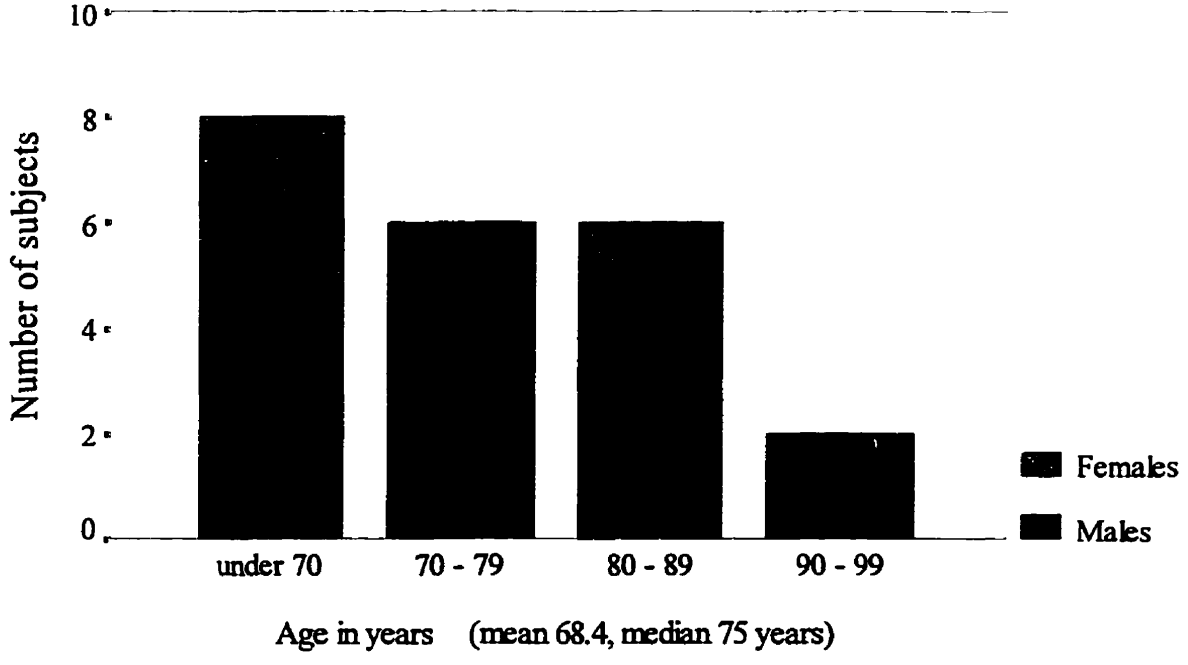


Figure 4.4c Randomization Phase subjects, Quick Response, by age group and sex

Reasons for presentation to the ER are shown in Table 4.2. Only the presenting complaint is listed here, but many of the older patients had underlying or co-existing problems: twenty-five subjects (31%) had hypertension or cardiac disease; 17 (21%) had diabetes mellitus, (6 were insulin-dependent); 8 subjects (10%) were demented and 6 (7%) had underlying malignancy. Twenty-eight people came to the ER following a fall: in 11 (39%) of these subjects this was accompanied by active or underlying medical problems.

Considering the acute problems which brought subjects to the ER, and the underlying and concurrent diagnoses, it appears that randomization worked out fairly well: the two groups were reasonably well matched, with the following exceptions: four individuals in the SC group had chronic neurological problems compared with two in the QR group; four (including one already listed under neurological problems) in the SC group were demented, but only three in the QR group; and two individuals in the QR group had carcinoma with metastases. Females were more likely than males to present with a fracture (7) or soft tissue injury (4) by comparison with 2 and 2. More males (6) than females (3) complained of lumbar or cervical spine problems.

Of the 81 patients who agreed to enter the study (with one subject in both phases), 21 did not complete all questionnaires. Table 4.3 lists the reasons for incomplete data for subjects in the two phases. Table 4.4 compares age, prior Barthel and initial questionnaire scores of those who completed initial and subsequent questionnaires, and those who for various reasons did not. Those who did not answer these questionnaires were somewhat older; available prior Barthel scores (reflecting function in daily activities before the incident which precipitated the ER visit) show that they were significantly less able than

Table 4.2. Reasons for presentation to the Emergency Room

Non-Random	Presenting Diagnoses	Randomized	
QR		SC	QR
16 (8 fractures)	**Orthopaedic**	11 (5 fractures)	13 (4 fractures)
2	**Neurological** (not back pain)	0	2
9	**Infection** skin and tissue	2	2
9	**Cardiorespiratory**	2	0
3	**Urological**	1	1
0	**Dementia**	2	1
0	**Diarrhea**	1	1
0	**Multiple Probs**	1	2
1	**Other**	0	0
n = 40		**n = 20**	**n = 22**

Underlying diseases included metastatic disease, dementia, Parkinson's disease, drug dependency and anxiety/depression

Table 4.3. Reasons for Incomplete Data

	n	Study Group	Data Available
Deaths	3	Pilot Phase QR	Prior and Initial Barthel questionnaires All Initial questionnaires All Initial and 2/52 questionnaires
Unable to complete due to illness	1	Pilot Phase QR	Initial Barthel questionnaire only
	1	SC	Barthel Initial, 2/52 & Final
	3	SC	All Barthel, Initial & 2/52 SIP & SF-36
	1	QR	All Barthel, Initial & 2/52 SIP & SF-36
Refusals	4	Pilot Phase QR	Prior and Initial Barthel questionnaires All Initial questionnaires Initial and 2/52 questionnaires All questionnaires exceot final SIP
	2	SC	Initial and 2/52 questionnaires All Barthel; Initial & 2/52 SIP & SF-36
	3	QR	Initial Barthel questionnaire Initial questionnaires Initial and 2/52 questionnaires
Lost to follow-up	1	Pilot Phase QR	All Initial and 2/52 questionnaires
	2	SC	

Where phase is not given, SC and QR denote randomized subjects

Table 4.4. Comparison of prior and initial scores of subjects who answered initial, second and final questionnaires with those subjects who for various reasons did not

	Answered initial, second & final questionnaires	**Did not answer all initial, second & final questionnaires**
n	61	21
Age (years) ± SD median, quartiles	68.2 ± 17.8 60.0, 74.0, 80.0	72.8 ± 20.2 71.0, 80.0, 85.0
Questionnaire scores median & quartiles		
Barthel prior to ER visit	n = 54 94.0, 100, 100	n = 16 84.0, 90.5, 100
Barthel at ER visit	n = 61 67.0, 84.0, 100	n = 20 44.5, 68.5, 89.5
Initial SF-36 Physical Function	0, 10.0, 40.0	n = 16 2.5, 7.5, 20.0
Initial SF-36 Mental Health	52.0, 68.0, 88.0	n = 16 36.0, 48.0, 68.0
Initial SIP Total	18.6, 25.7, 33.4	n = 16 27.4, 33.9, 38.0
Initial SIP Physical Dimension	18.2, 29.1, 43.9	n = 16 26.2, 37.5, 44.3

those whose data at the three time points were complete (Wilcoxon rank sum test $p = .036$); initial (ER) Barthel scores also demonstrate their poorer physical ability (Wilcoxon rank sum test $p = .022$).

Nine subjects who initially agreed to take part subsequently declined to answer questionnaires: they were older (80.3 $\pm$ 6.7 years, quartiles 79.0, 81.0 and 83.0 years) but their prior function was almost as good as those whose data was complete: (prior Barthel score 94.6 $\pm$ 6.6%, median 99%). Their initial function and health scores were essentially the same as others in the group with incomplete data. It is possible that these subjects were daunted by the time and effort required to complete questionnaires (at least one hour, and often longer.) The age and scores of these nine subjects are included in the data for the 21 subjects who failed to complete all questionnaires (Table 4.8).

The Barthel questionnaire was brief, objective and easy to answer.
In the Randomization Phase, while complete Barthel data were obtained for 17/20 SC and 19/22 QR subjects, complete SIP and SF-36 data could be obtained only for 12 SC and 18 QR subjects. Comparison of available Barthel data shows that of those in the SC group, subjects with incomplete data fared worse at all time points. Missing Barthel data at 3 months in randomized subjects was due to refusals (4) and failure to find subjects at follow-up (2). Available information about these subjects suggests that they all progressed satisfactorily during the study timeframe, so it is not unreasonable to consider the best possible scenario in terms of the 3-month Barthel score, *i.e.* to substitute 100% Barthel scores for these 6 subjects, and on this basis to calculate group means. Even when this was done, the "best possible" 3-month Barthel score for the SC "data incomplete" group

was 58.5 ± 36.2%, median 47.0%, much less than the score for "data complete" SC subjects, see Table 4.5. There was a smaller difference in corresponding data for QR subjects, "best possible" data being 75.0 ± 50.0%, median 100%. This is the dilemma in reporting scores: not including data belonging to subjects whose data are incomplete in this study gives an unjustifiably rosy picture of subjects at all stages; omission of data introduces an important bias because these subjects are not represented, and scores will only give the picture for subjects with better health. On the other hand, reporting all available data makes interpretation difficult, because the initial data are loaded with low scores representing people who for various reasons dropped out, and 3-month scores unless regarded with care might suggest greater mean improvement than actually occurred.

Figures 4.5 to 4.9 include all available data, recognizing that the SF-36 and SIP 2-week and particularly the 3-month data do not include some subjects whose health was worse than the mean. In the Randomized Phase, of the 8 SC subjects for whom full questionnaire data were not obtained, 4 were too sick to answer and another was in poor health; in the QR group, of the 4 subjects with incomplete questionnaire data, 1 was too sick (and died shortly after his 3 months in the study). Three-month Barthel questionnaireswere completed with assistance from relatives or caregivers for these individuals. Of the 6 subjects not accounted for, 2 are known to have returned to work; Community Health records and the MCP record of physician services show only minor items for the other 4 subjects. Thus while final (3-month) Barthel data are weighted slightly pessimistically (because they do not include these 6 subjects), final SF-36 and SIP data are biased in the other direction because data are lacking for 12 subjects who as a group fared worse.

Table 4.5: modified Barthel data, Randomized Phase: quartile scores (%) prior to the ER visit, at the time of the visit & at 2 weeks & 3 months.
Subjects who answered all initial and subsequent questionnaires compared with those who for various reasons did not.

	Standard Care n = 20		**Quick Response Care** n = 22	
	Data complete n = 12	Data incomplete n = 8	Data complete n = 18	Data incomplete n = 4
Prior Barthel	86.0, 100, 100	54.0, 90.5, 100	88.0, 97.5, 100	n = 3 90.0, 100, 100
ER Barthel	65.5, 74.5, 100	37.0, 76.0, 89.5	49.0, 72.5, 93.0	n = 4 36.0, 75.5, 99.5
Two week Barthel	63.0, 89.0, 100	26.0, 76.5, 100	80.0, 97.0, 100	n = 2 10.0, 50.0, 90.0
Three month Barthel	89.5, 100, 100	n = 5 20.0, 34.0, 37.0	85.0, 94.5, 100	n = 1 0

For a variety of reasons, 32 potential subjects referred to the QR study were not enrolled: one was medically unsuitable; the needs of 11 could be met by routine Community Health services; 4 lived outside the Community Health (St. John's) area; 7 preferred not to take part and family members refused on behalf of 4 others; and 5 were unable to comprehend what the study involved). Although the mean age of potential subjects not enrolled was 5 years greater than that of those who were enrolled, this difference was not significant (Student t-test $p = 0.69$). Available demographic data for these subjects given in Table 4.6 suggest that in most respects they were similar to those who took part.

In terms of diagnosis, the only difference was that while nearly half of those who were enrolled had orthopedic problems (40/82: 17 fractures, 10 soft tissue injuries, 13 back or neck injuries), only one-fifth (6/32) of non-participants had orthopedic problems (5 fractures, 1 back injury). Other than this, although 28% of patients referred to the QR project did not take part, enrolled subjects appear to be reasonably representative of the target population of adults who may benefit from a brief QR intervention.

Table 4.6. Demographic characteristics of enrolled subjects compared with those of potential study subjects not enrolled

Characteristic	Enrolled	Not enrolled
n	82	32
Age (years) ± SD quartiles	69.4 ± 18.5 63.0, 75.5, 81.0	n = 29 74.8 ± 10.0 67.0, 78.0, 81.0
Percentage female	56.0%	53.0%
Married **Single** **Widowed** **Divorced/Separated** **Unknown**	35.4% 22.0% 35.4% 7.3%	37.5% 3.1% 34.4% 6.3% 18.8%
Family/friends support network	85.4% positive	81.3% known positive

4.1 Health Status as measured by Questionnaires: unless otherwise stated, comparisons will be drawn between the randomized SC and QR groups only.

Based on reports in the literature where clear clinical differences are anchored to differences in questionnaire scores, (see pages 30 - 34), the following group mean differences will be taken as evidence of a real clinical difference:

modified Barthel Index, 7%; SF-36 Physical Function, 13%,

SF-36 Mental Health 30%; SIP Total score, 7% and SIP Physical Dimension, 10%.

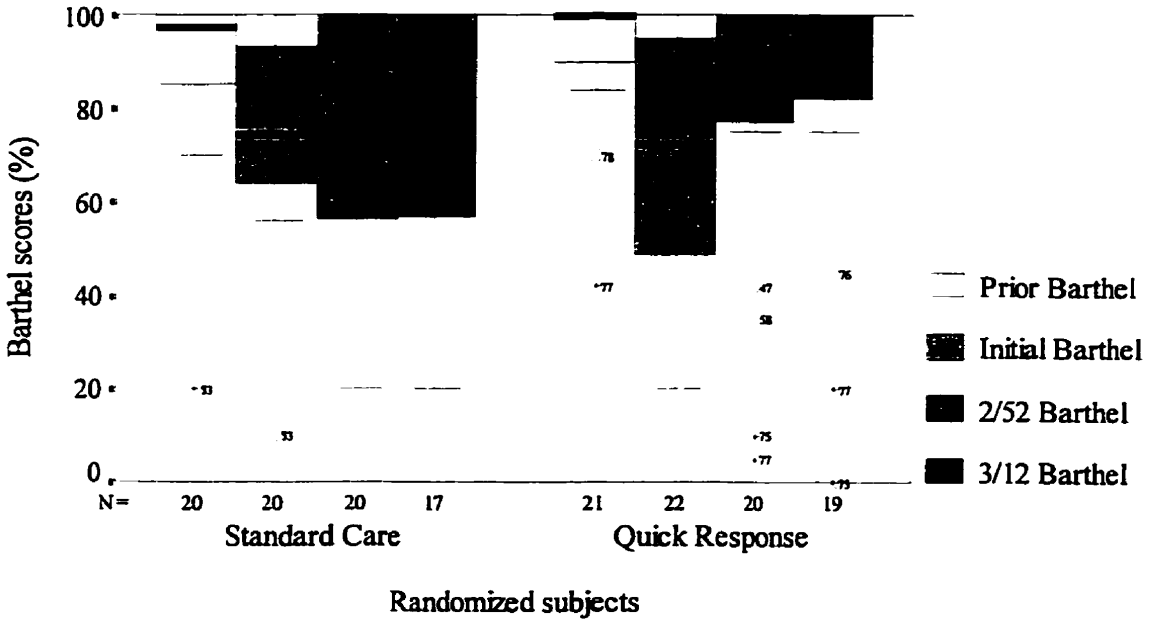


Figure 4.5 Modified Barthel scores: box and whisker plot showing median, 25th and 75th percentiles

There was no difference (clinical or statistical) between the randomized SC and QR groups in baseline Barthel scores (reflecting function in ADL prior to the ER visit), Wilcoxon rank sum test, p = 0.31, or in initial scores : Barthel p = 0.29; SF-36 Physical Function p = 0.38; Mental Health p = 0.41; SIP Total p = 0.49, SIP Physical Dimension p = 0.44, (Figures 4.5 - 4.9).

Box and whisker plot definitions are given in the Appendix, p. 128.

At two weeks, the Barthel scores show that in terms of ADL, although there was no significant clinical or statistical difference between the groups (Wilcoxon rank sum test p = 0.34), by comparison with their ER status QR subjects had improved significantly (Wilcoxon signed ranks test p = 0.01 whereas improvement in the SC group was not significant clinically or statistically (signed ranks test p = 0.07).

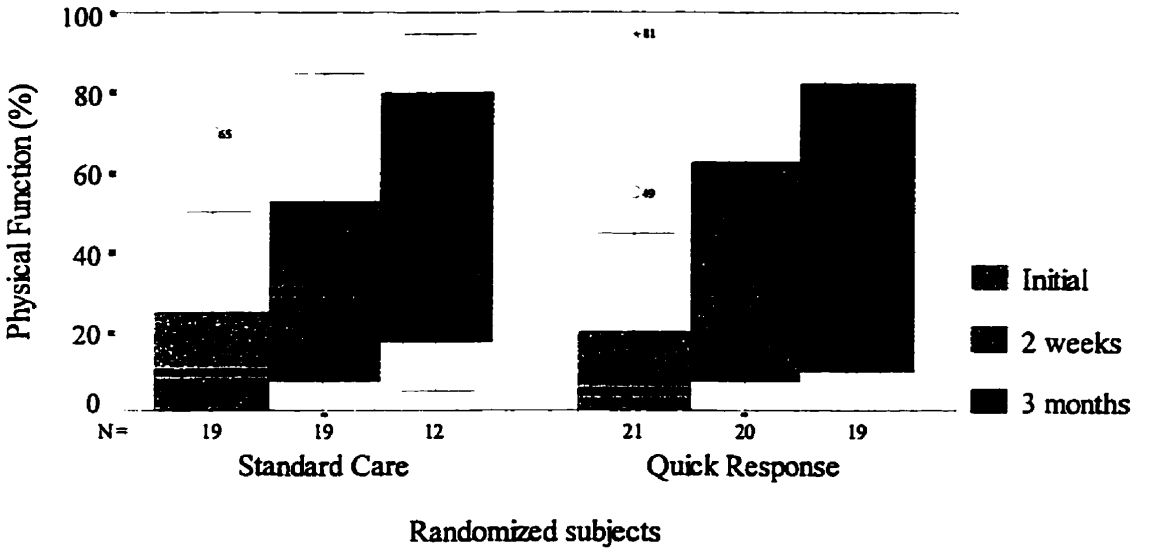


Figure 4.6 SF-36 Physical Function: box and whisker plot showing median, 25th and 75th percentile scores

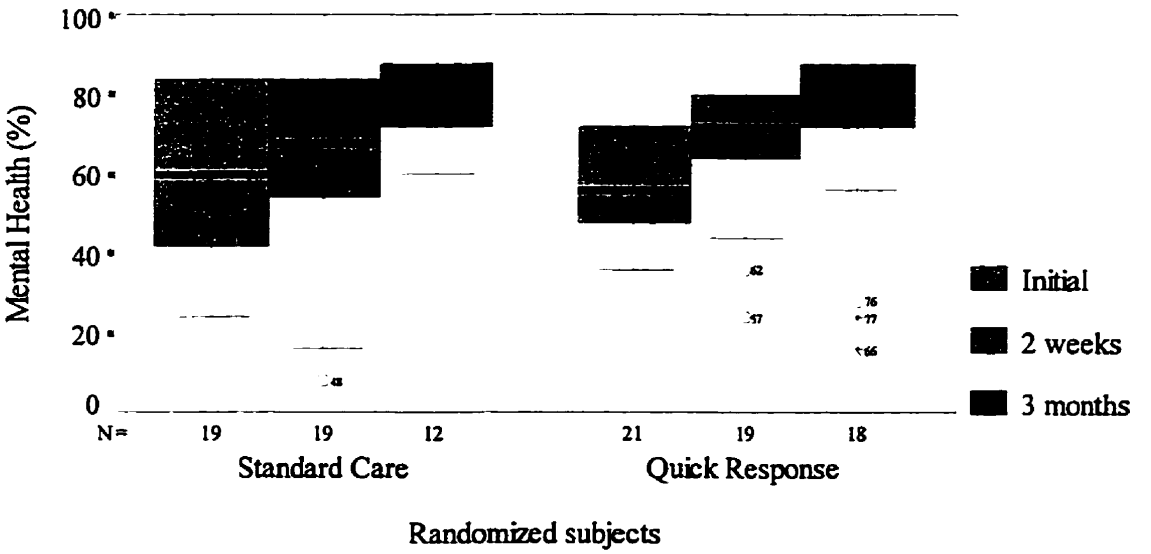


Figure 4.7 SF-36 Mental Health: box and whisker plot showing median, 25th and 75th percentile scores

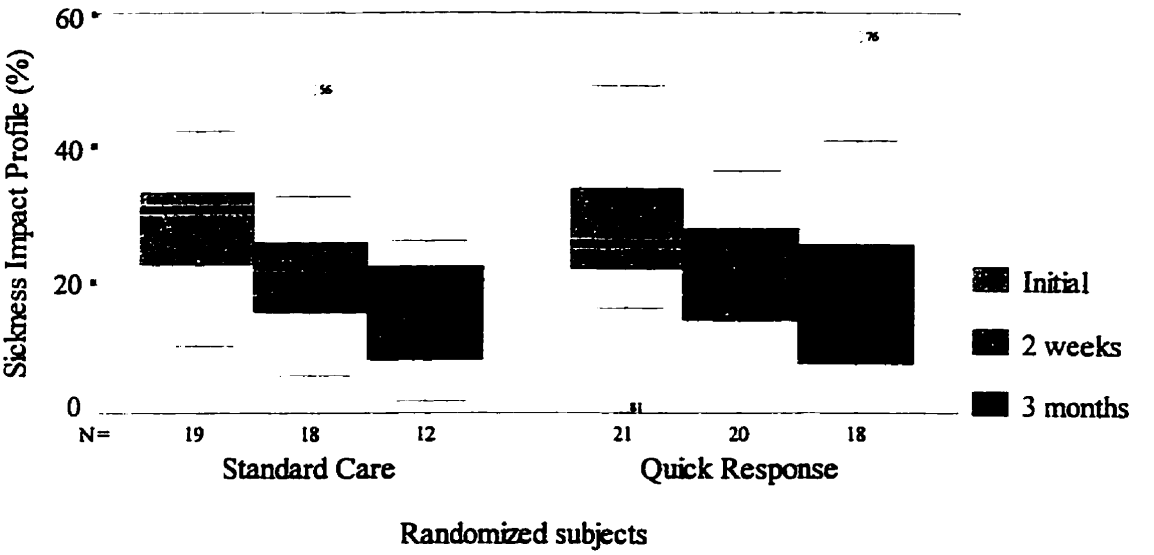


Figure 4.8 Sickness Impact Profile: box and whisker plot showing median, 25th and 75th percentile scores

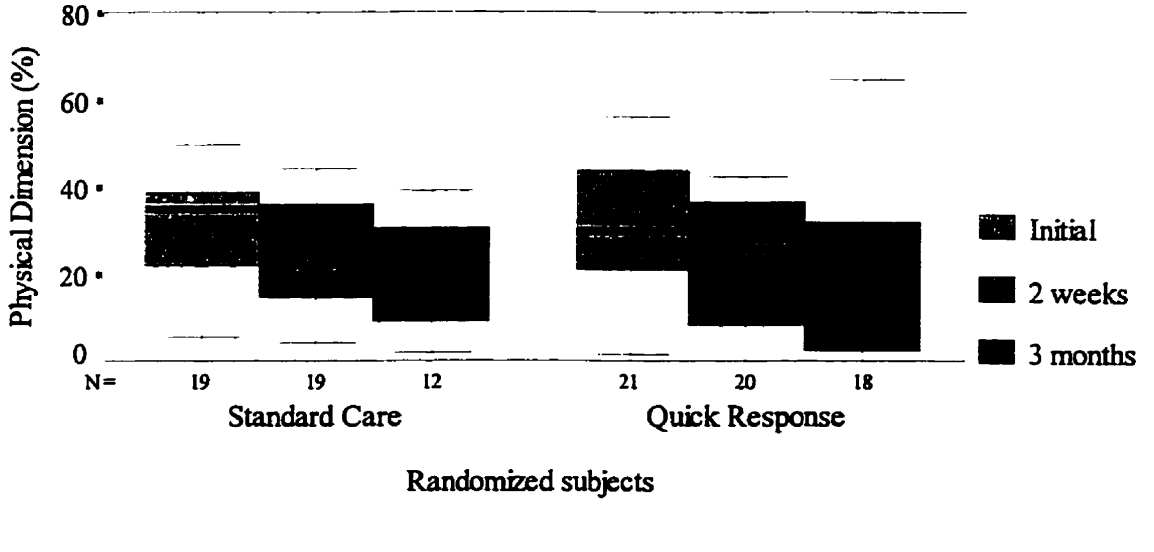


Figure 4.9 Sickness Impact Profile Physical Dimension: box and whisker plot showing median, 25th and 75th percentile scores

With the exception of the SF-36 MH domain, where improvement noted at two weeks was insignificant both clinically and statistically (SC $p = .11$, QR $p = .20$), other questionnaire data showed marked significant improvement in both groups at two weeks by comparison with their initial scores: SF-36 PF, $p < .001$, Total SIP, $p \leq .005$ and SIP PD, $p \leq .01$, (Wilcoxon signed rank sum test).

As expected, gradual improvement in physical function continued in both groups and Barthel and SF-36 (PF) and SIP scores at three months demonstrate this. Two weeks, while adequate to regain a measure of independence, was shorter than the time required for full recovery from many of the conditions seen in the ER: *e.g.* fractures or back injuries.

At three months, no significant differences between the randomized SC and QR groups were seen: Barthel Index, $p = .39$ (Figure 4.5), the SF-36 Mental Health domain, $p = .48$ (Figure 4.7), or SIP Total, $p = .36$ or Physical Dimension, $p = .40$ (Figures 4.8, 4.9). The differences (SC higher than QR) in SF-36 Physical Function scores, mean (10.8) and median (37.5), were clinically but not statistically significant, $p = .20$, (Figure 4.6); it is worth noting that 6 subjects (5 in the SC group, and one in the QR group) represented in Barthel data were not included in SF-36 or SIP data. These 6 subjects are known to have had poor health outcomes. Six other subjects, three in each group, are also not represented in the three month scores (Barthel, SF-36 or SIP) but as far as is known from Community Health and MCP records, their health care needs were minor during the three-month period.

Although differences observed between the groups were not statistically significant, numbers in the two randomized groups were much too small to be able to demonstrate that there were no real differences. Freiman *et al.*[45] discussed how small sample sizes often led investigators to conclude that there was no therapeutic effect in an intervention when their inadequate sample size did not provide adequate power to be able to reject the null hypothesis that there was no treatment effect. Julious *et al.*[46] discussed sample sizes needed to give a reasonable chance (power) of detecting a predetermined difference in an outcome measure such as the various dimensions of the SF-36. Assuming significance of 5% and power of 80%, he compared the numbers needed in each group using parametric and non-parametric methods, and showed that for the Physical Function dimension, parametric techniques dictated a sample size of 285 subjects, and non-parametric, 544 subjects to detect one discrete value (in this case, 5%) below the population mean, and 247 to detect one discrete value (again 5%) above the mean of a population of patients with relatively minor medical conditions. The corresponding sample sizes needed to detect a difference of one discrete value (4%) in the Mental Health dimension were 358 (parametric), and 738 and 217 (non-parametric calculations). Parametric calculations assume that data have a normal distribution, but where scores are skewed, non-parametric calculations are more appropriate.

Sample size calculations in this study were made on the basis of detecting a reduction in costs of 25% in the intervention (QR) group by comparison with the SC group, with significance level 5% and power 80%; the calculated sample size in each group was 315 subjects. Had study enrolment reached the planned number, numbers

would have been sufficient by Julious' calculations (using the clinically significant changes for SF-36 PF and MH identified on page 53, and the standard deviations we found) to provide sufficient power to detect differences in SF-36 scores if they had occurred.

Table 4.7 shows patient admissions to hospital and length of stay. This information was obtained from hospital data bases, and confirmed by referring to MCP records of physician services. Three subjects were admitted to hospital for a combined total of 21 days as part of their initial Standard Care. Another 4 subjects in the SC group could not manage at home and had to be admitted to hospital during the first two weeks following their ER visit.

By comparison, no randomized QR subject was admitted to an acute or chronic hospital bed in the first two weeks. When non-randomized QR subjects are included, the 61 QR subjects with a total of 108 days in hospital (of which 80 were acute) had 58% fewer hospital days per subject than the SC group of 20 subjects with 85 days (34 acute), ($p = .04$, Wilcoxon rank sum test).

Over the three-month period (including the first two weeks) 22 randomized QR subjects spent 100 days in hospital (40 acute), 50.5% fewer per subject than the 184 days (133 acute) spent by SC subjects ($p < .0001$, Wilcoxon rank sum test). Including non-randomized QR subjects, the QR group spent 380 days in hospital (242 acute), 31.5% fewer days per subject than the SC group ($p = .09$).

Table 4.7 Patient Admissions to Hospital and Length of Stay by Study Group

	Acute Care		Chronic Care		Palliative Care	
	Admission date		Admission date		Admission Date	
Study Group	**1 - 14 days**	**15 - 91 days**	**1 - 14 days**	**15 - 91 days**	**1 - 14 days**	**15 - 91 days**
Non-Randomized QR n = 40 group mean days	5 admissions 80 days 2.00	7 admissions (5 subjects) 122 days 3.05	1 admission 28 days 0.70	1 admission 50 days 1.25		
Randomized Standard n = 20 group mean days	4 admissions 34 days 1.70	2 admissions 99 days 4.95	3 admissions 51 days 2.55			
Randomized QR n = 22 group mean days		5 admissions 40 days 1.82				1 admission 60 days 2.73

Admission date "1 - 14 days" refers to hospital admission during the first 14 days after the ER visit (the timeframe of maximum QR intervention), and "15 - 91 days" to admission in the 11 weeks thereafter. Length of stay refers to days of care within the QR 3 month timeframe.

Hospital care, particularly in acute care beds, is costly and the difference between SC and QR subjects in terms of hospital days is reflected in their hospital costs, (including cost of hospital visits and out-patient procedures), Table 4.8 and Figure 4.10. The mean (median) cost for each of 20 SC subjects was 1.9 (1.8) times that for each of the 22 randomized QR subjects, Wilcoxon rank sum test $p > 0.2$. On the other hand, the mean (median) Community Health cost per SC subject was only 0.5 (0.6) that per QR subject, Wilcoxon rank sum test $p > 0.1$. Total costs include all publicly-funded expenditures, not only hospital and Community Health services, but also the cost of long-term care in institutions and personal care homes (for those who could afford to contribute), physicians' billing through MCP, and homemaking covered by the Department of Social Services, and homemaking and nursing services funded by the Department of Veterans' Administration. Overall, for the 3-month period following the ER visit, the mean (median) cost of all care for SC subjects was 1.4 (1.4) times greater than for QR subjects. These differences were not statistically significant (Wilcoxon rank sum test, $p > 0.5$). The same trends are evident when all QR subjects (non-randomized and randomized) are compared with the SC subjects.

Total health care costs over 3 months were less than $1,000 for one-third of study subjects (27/81), and less than $10,000 for 86% (70/81). For subjects whose 3-month costs were less than $10,000 there was no difference between SC and QR. The observed difference between the two groups lay in the the 11 subjects whose care costs exceeded $10,000. The mean (median) cost of the 3 SC subjects was $27,408 (27442), and of the 8 QR subjects, $18,101 ($15456), $p = .06$. Only 6 subjects, (3 in each group, *i.e.* 15% of the SC group but only 4.9% of the QR group) had 3-month costs exceeding $20,000.

Table 4.8. Comparison of Costs over 3 months between Standard and Quick Response groups: Community Health, Hospital and Total Costs

Group	**Community Health Costs ($)** mean ± SD quartiles	**Hospital Costs ($)** mean ± SD quartiles	**Total Costs ($)** mean ± SD quartiles
Standard n = 20	761 ± 1273 95, 336, 579	4102 ± 7977 241, 612, 4097	6292 ± 9457 904, 2824, 5194
Quick response n = 22	1425 ± 2660 230, 522, 905	2108 ± 4543 155, 345, 2043	4528 ± 6522 585, 2063, 4882
Non-randomized **Quick response** n = 40	1071 ± 1333 183, 429, 1612	2848 ± 5494 241, 285, 3495	4652 ± 6353 678, 1819, 5907

Total costs include homemaking services funded by DVA (Department of Veterans' Affairs) and DOSS (Department of Social Services), long-term care costs and physicians' billing as well as Community Health and hospital costs.

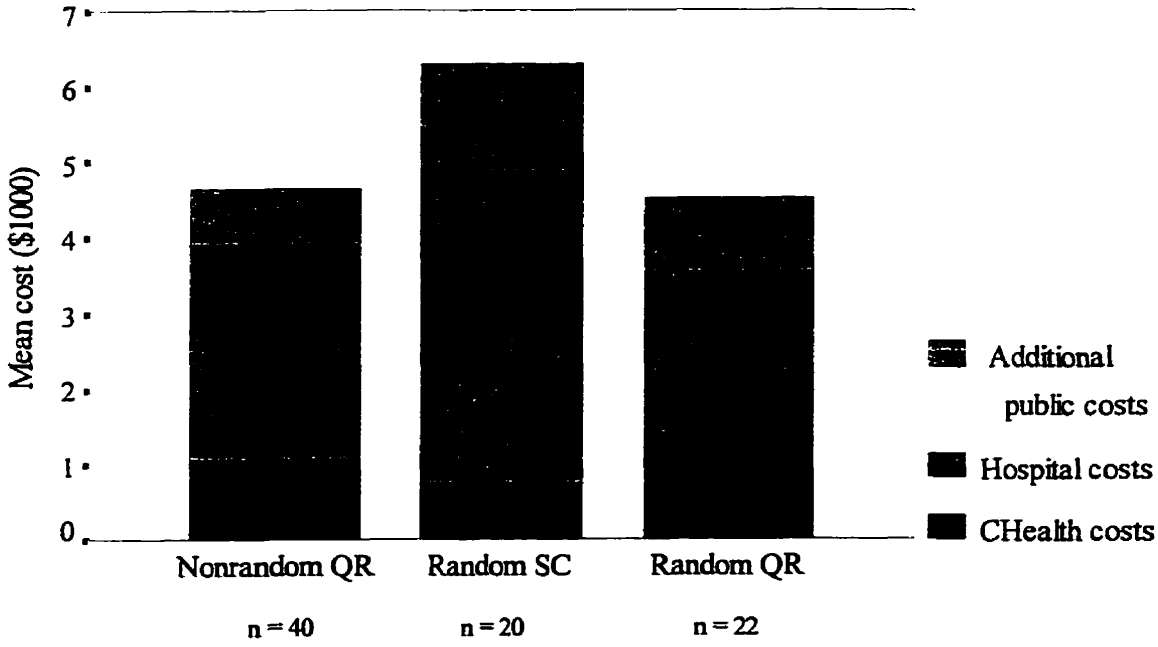


Figure 4.10 Mean cost of all services provided over 3 months, by group: Community Health, hospital and other publicly-funded services

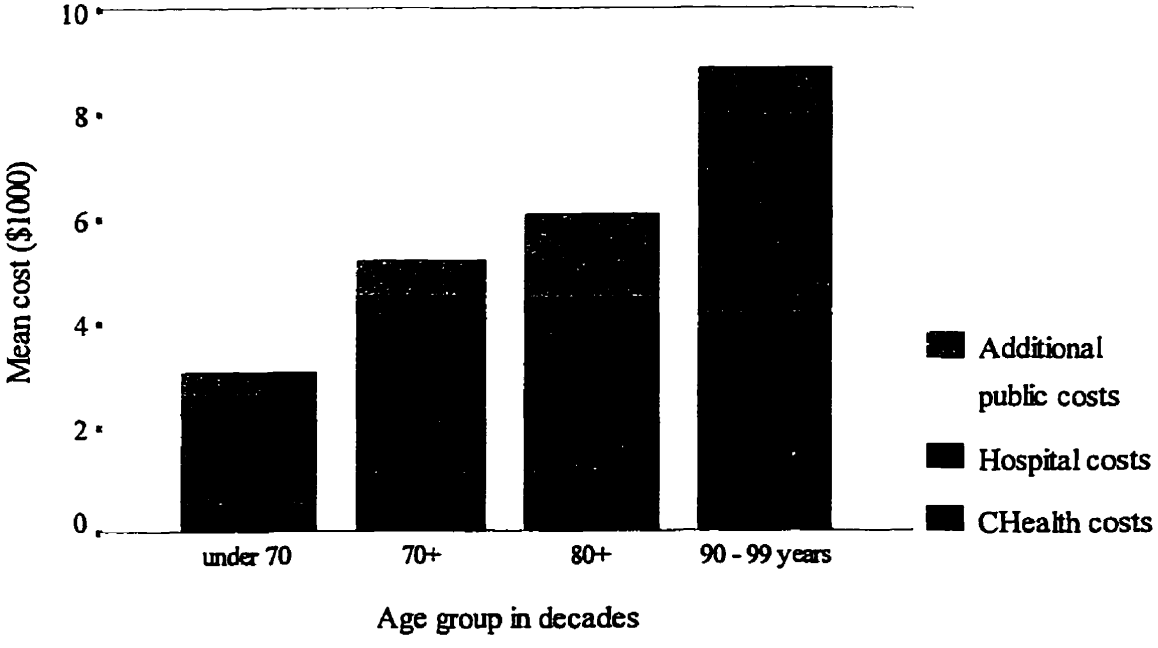


Figure 4.11 Mean cost of services provided over 3 months, by age group: Community Health, hospital and other publicly-funded services

Total costs include homemaking services funded by DVA (Department of Veterans' Affairs) and DOSS (Department of Social Services), and physicians' billing, as well as Community Health, hospital and long-term care costs.

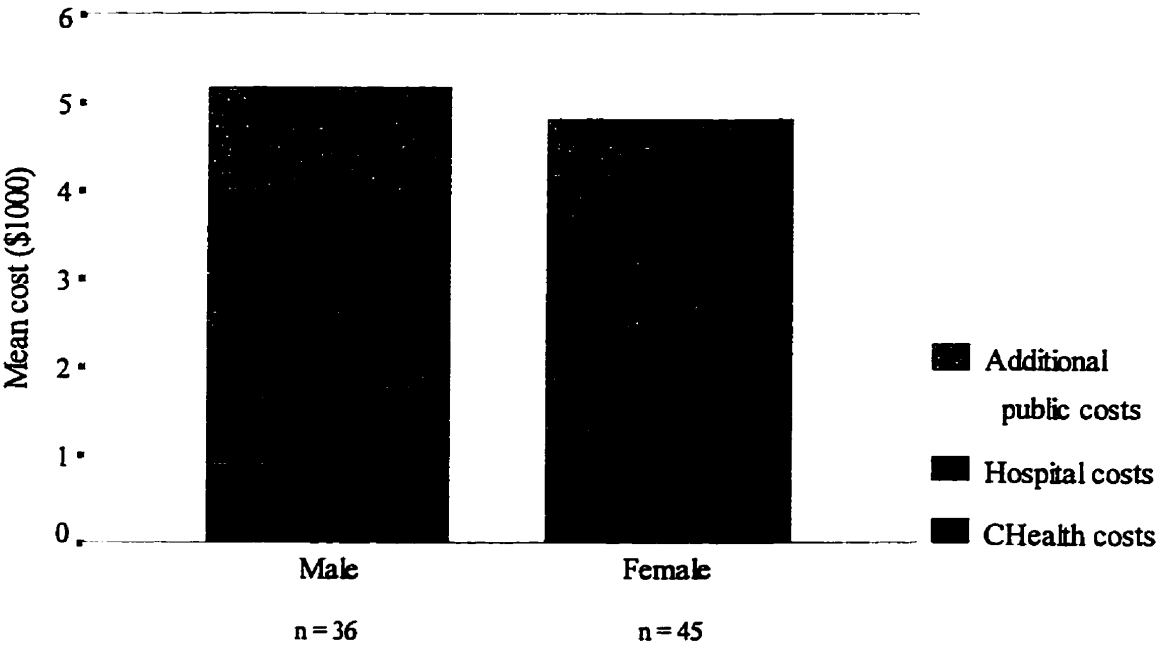


Figure 4.12 Mean cost of all services provided over 3 months, by sex: Community Health, hospital and other publicly-funded services

Total costs include homemaking services funded by DVA (Department of Veterans' Affairs) and DOSS (Department of Social Services), and physicians' billing as well as Community Health, hospital and long-term care costs.

Data displayed in Figure 4.11 show (as expected) that the overall cost of care increases with age. There was no difference between the sexes in terms of total costs over 3 months, Wilcoxon Rank Sum test, $p = 0.48$ (Figure 4.12). Community Health mean (median) costs were higher for females than males by a factor of 1.4 (1.7), $p = .06$.

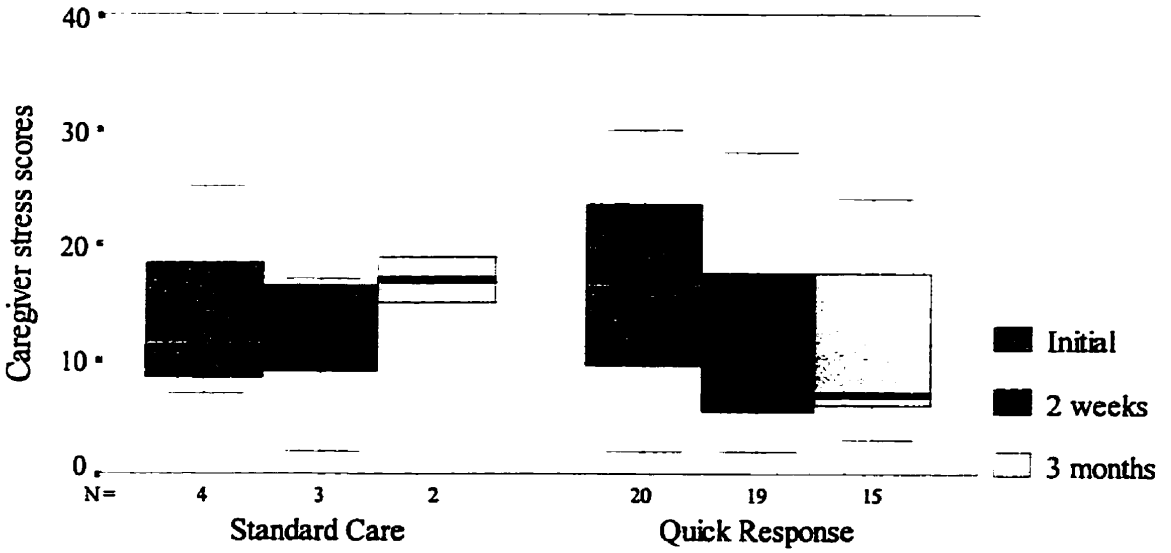
4.2 Impact on the family: thus far, family costs have not been considered. These are more difficult to quantify, in part because many families regard it as a natural part of life to care for a family member who falls sick, and they do not feel it necessary, or appropriate, to count the cost, whether personal or financial. Only when their burden becomes intolerable do they sit down and calculate the cost (and then with a considerable sense of guilt). Subjects and/or family members were asked to report in a daily log the extra costs associated with being sick or with caring for a sick relative, in terms of time, dollars and disruption of routine activities or loss of leisure time. Several family caregivers had difficulty in recalling this information, or perhaps felt it was not proper to itemize the costs in such a clear fashion, but with some encouragement would allow that their social life had suffered to some extent. Some respondents were much more articulate when responding to these questions, because they had already given some thought to the subject, and in some cases because they perceived it was important for their own health that they not get overwhelmed. Twenty-five subjects identified family caregivers, 14 in the non-randomized phase, and 12 (4 SC, 8 QR) in the randomized phase (with one represented in both phases). Twenty caregivers (80%) were female: 11 were wives, 3 were sisters and 3 were daughters. Of the 5 male caregivers, 1 was a husband, the others were sons. Twelve people indicated that caregiving was a full-time job, leaving them no leisure time unless they arranged - and paid for - alternative care. Seventeen (68%) of the 25 subjects who named caregivers lived in houses owned by themselves or their children; 9 of these subjects carried health insurance. In all, 12 (48%) of the subjects with caregivers had health insurance; at least 6 subjects (24%) were receiving the

Guaranteed Income Supplement, which also entitled them to purchase their medications more cheaply, using a drug card. Ten of the 25 caregivers could not identify any additional dollar costs connected with caregiving, others listed costs ranging from $15.00 for equipment such as a raised toilet seat and $20.00 per week for Attends to $4500 for supervision and care 5 days a week for 5 weeks. Median cost for families caring for an SC patient was much greater over the 3 months than for families whose relative received QR: $2960 compared with $55, $p < .01$ (Wilcoxon rank sum test). This might reflect an underlying difference in financial status between families of the SC and QR groups. The range of support provided to a sick person and his or her family varied widely, usually depending on ability to pay for services. Families with incomes above a certain level have to pay part of the cost of care in long-term institutions, as well as the cost of homemaking services (with the exception of homemaking services provided as part of Quick Response). One caregiver noted with some anger that a private organization providing home support services charges private individuals more than twice the hourly rate it pays its employees. She suggested the provincial government might pay family caregivers a small allowance when they devote a major part of their time to caring for a sick person in the family. Until recently this informal contribution to health care has been largely unmeasured and almost wholly unpaid. Governments have been reluctant to recognize, even in a small way, the hours of care provided by some families to keep handicapped or elderly relatives at home. Five caregivers were employed outside the home, and three recorded that they had lost time from work. Another one would have lost time from work had the crisis not occurred during school holidays. The other caregivers were either retired or not employed.

4.3 Measurement of stress and the health of caregivers: stress experienced by caregivers was measured using the RSS with the 3-point scoring system (*i.e.* maximum possible score for the 15 questions would be 30) devised by Eagles *et al.*[40]. These authors demonstrated a clear difference, as measured by the RSS, in the burden felt by family caregivers depending on whether the elderly relatives for whom they were caring were demented or not (median scores 5.75 and 1.71, $p = 0.001$). As noted on pages 28 and 29, with an ordinal scale like the RSS it is more appropriate to consider group medians rather than the mean scores. RSS scores are therefore reported in terms of the median, with the clinically significant score difference taken to be 4 points.

The GHQ-30 was designed to be a screen for the presence or absence of psychiatric morbidity, and on this basis individual score changes to/from negative (≤ 4) from/to positive (≥ 5) are taken to be clinically significant.

Twenty-four subjects had relatives who answered the Relatives' Stress Scale (RSS) and the General Health Questionnaire (GHQ-30), but information at all time points was available only for 16 of them. Initial scores for the 9 who did not complete questionnaires at all time points revealed no difference in terms of the RSS (Wilcoxon rank sum test, $p > .05$) or the GHQ-30 (χ^2 test, $p > 0.5$) between them and family members who completed all RSS and GHQ-30 questionnaires. Figures 4.13 and 4.14 show RSS and GHQ-30 scores for family members of subjects in the SC and QR groups.

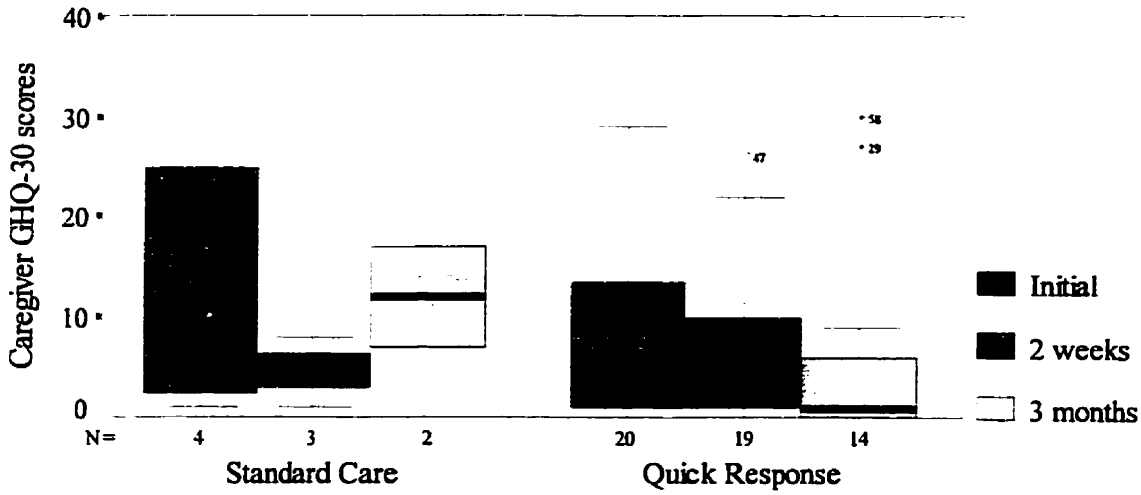


Lower scores reflect less reported stress

Figure 4.13 Relatives' Stress scale; box and whisker plot showing median, 25th and 75th percentile scores

There was a statistically significant baseline difference of 12 points in the initial RSS median score between carers of randomized SC and QR recipients ($p = 0.01$, Wilcoxon rank sum test); when carers of non-randomized QR subjects were included, the clinically significant baseline difference of 5 points in the median RSS score was not statistically significant ($p = .30$). The difference in perceived caregiver burden was not reflected in the initial GHQ-30 status, where there was no difference between the two groups (χ^2, $p > 0.5$).

At two weeks, caregivers of Standard Care subjects reported more stress, while QR caregivers reported minimal improvement.



Scores above 4 suggest psychiatric distress

Figure 4.14 General Health Questionnaire-30: box and whisker plot showing median, 25th and 75th percentile scores

By three months, three SC caregivers continued to report higher levels of stress than previously (although scores are only available for two). Of the caregivers not represented (half the group), one dependent relative progressed well and the other continued in very poor health: the family had major difficulties in coping until he was placed in long-term care after the end of the study timeframe. By contrast, the distress felt by the QR group of caregivers had diminished. Of the two caregivers not represented, one sick person had died, and the other had been placed in long-term care, to the satisfaction of everyone concerned. The median RSS score for caregivers of QR subjects was 5 points lower than the two week median score, but the difference was not statistically significant ($p = .24$).

Table 4.9 Psychiatric distress in caregivers: General Health Questionnaire-30

	Non-random n = 14		Random Standard Care n = 4		Random Quick Response n = 8	
Status	+	-	+	-	+	-
Initial % caregivers	9 (64%)	4 (29%)	2 (50%)	2 (50%)	6 (75%)	2 (25%)
2 weeks % caregivers	6 (43%)	6 (43%)	2 (50%)	1 (25%)	5 (62.5%)	3 (37.5%)
3 months % caregivers	3 (21%)	6 (43%)	2 (50%)		2 (25%)	4 (50%)

Negative: GHQ score ≤ 4; Positive: GHQ score ≥ 5

Missing data were due to refusals (5); regular caregiver away to get a break (1); relative placed in long-term care out of town, family expressed satisfaction with the arrangement, but did not complete questionnaire (1); relative deceased, caregiver was not asked to complete the questionnaire (2).

Table 4.9 expresses the GHQ data in terms of psychiatric distress status as defined by Goldberg: ≤ 4, negative; ≥ 5, positive. Overall, caregivers' RSS and GHQ scores improved between the ER visit and the 3-month follow-up, but 6 caregivers (25% of the caregiver sample) scored > 15 points on the RSS, and 7 had GHQ scores > 5, *i.e.* positive for psychiatric morbidity at 3 months. This illustrates the very considerable burden carried by caregivers. It is not possible to draw any conclusion as to whether provision of QR support made any real difference in terms of stress relief or the health of caregivers, firstly because numbers in the randomized phase were inadequate, particularly in the SC group, and secondly, because it appears there was an initial baseline inequality (as measured by the RSS) in the caregivers of the two randomized groups. Personal observation suggests that although support with nursing care and assistance with homemaking alleviates some of the practical problems, it does not lessen the distress of watching a loved one deteriorate.

4.4 Satisfaction with care: the Satisfaction questionnaires (patient and caregiver) were developed for this project, and therefore have not yet been validated. At this stage, it is not clear what represents a clinically significant score difference. Thirty-three subjects and 12 caregivers answered the satisfaction questionnaires two weeks after enrolment. The median subject score was 84 (out of a possible 98 representing "delighted" in answer to all 14 questions), indicating a high degree of satisfaction with the community care provided. Few patients or caregivers were willing to criticize the services they were receiving, mainly because they appreciate the care, and it is not part of their culture to express dissatisfaction in this situation. Even when they were encouraged to speak out,

few offered any critical comments. It is possible that they perceived the research assistant to be part of the health care system, although the role of health research as distinct from care was explained.

There was a statistically significant difference in patient satisfaction scores between the SC and QR groups, Wilcoxon p = .001, see Table 4.10, and this difference was borne out by a number of comments made by subjects in each group. Scores should be interpreted with some reservation because the number of respondents was low, particularly for the caregiver questionnaire, and scores may not be representative of the whole sample. Furthermore, SC subjects knew that randomization had assigned them to the non-intervention group in terms of QR, and therefore assumed (correctly) that other study subjects might have received care more promptly, with more homemaking, at no extra cost, and perhaps more frequent professional visits as well. Even so, there were only a few real (or justified) criticisms with the arrangements made in the two weeks following the ER visit, and of those, none concerned the care provided by Community Health.

Table 4.10 Patient satisfaction with community care given them in the two weeks following their ER visit

	Standard Care n = 11	**Quick Response** n = 22
Patient satisfaction score quartiles	68.6, 79.8, 82.6	82.6, 86.1, 93.8

The Caregiver Satisfaction Questionnaire was answered by 12 respondents, of whom 9 were caregivers of QR recipients. The median score was 96.6 in the SC group and 84.7 in the QR group. No question received a score lower than 4 (the neutral point in the Delighted-Terrible scale). Caregivers of the QR group were more critical of services provided than were their sick relatives. Caregivers of SC recipients commented that although they had some criticisms of the system, the homemakers and professional staff who came were excellent.

The questionnaire dealt largely with the skill and personal characteristics of care providers, and only one question addressed cost, so some of the larger concerns of families caring for very dependent relatives, such as recognizing the contribution of families providing essentially 24-hour care, were not reflected in the overall score.

There were many comments concerning the length of time subjects had to wait in the ER before being examined by ER staff, but it was not part of the study mandate to do more than acknowledge this and express sympathy. Participation in the QR study added to the time spent in the ER (usually about one hour and twenty minutes on top of about three hours before the QR nurse was paged), and some of the patients and/or their caregivers who did not wish to take part in the study may have become exasperated by the long wait.

4.5 Inadequate enrolment of subjects. The feasibility study reported that in 1992 in a one-month period at the three city ERs, 75 patients, *i.e.* 11% of those who met age and residency criteria, would have been suitable for QR services had they existed. In the 1995 QR study, only 113 patients were referred and 81 (1 twice) were enrolled, one-fifth of those expected over the 5½ months of QR. Of these, only 16* (14.2%) either would have been, or actually were (as part of their Standard Care) admitted.

The original premise of the QR project was that unnecessary admissions might be avoided, but between the time of the feasibility study and implementation of the QR project, the number of acute care beds in St. John's had fallen from 857 to 742, a 13% reduction. This had occurred in response to a recommendation of the Royal Commission on Nursing Homes and Hospital Costs (1984). One of the consequences was a change in patterns of admission to hospital: many people who would have been admitted in 1992 were less likely to be candidates for admission in 1995. During the same time period, visits to an ER rose by ~ 5.4%. On this basis although one might have expected no difficulty in 1995 in recruiting subjects who could benefit from extra support at home, the number of subjects referred by ER physicians was much lower than predicted.

* Nine people during the Pilot Phase, and 2 during the Randomization Phase would (in the absence of QR) have been admitted to hospital; 3 people were admitted during the Randomization Phase as part of Standard Care, and 2 people who refused to enter the study were also admitted to hospital.

One possible reason for poor enrolment in the study was the climate of uncertainty and apprehension prevailing in hospitals at the time. Reduction of transfer payments from Ottawa to the provinces started in the 1980s and forced the cash-strapped Newfoundland government to evaluate the organization and cost of health care delivery. Staff in 1995 knew that the general plan was to bring the hospitals under one management, to rationalize the Emergency Departments, to close hospital beds and to provide more care in the community. What they did not know was how or when this plan would be translated into action and how many (and whose) jobs would be affected.

The team of QR nurses was selected from Community Health staff; they operated in a rota in the city ER departments, working in "hospital territory". In the atmosphere of tension surrounding the future of hospitals and their ER departments, it is possible that enthusiasm for the QR project was sometimes lacking. ER staff work under considerable pressure, and the presence of an "outside" nurse who could not assist with the regular work load may not have been perceived in a positive light.

Another factor was a reluctance on the part of some emergency physicians and consultants to allow sicker patients to be cared for at home when they were not certain how good medical coverage would be if for whatever reason, the patient's condition worsened. Randomization with its inherent uncertainties as to patient management may have made some referring physicians less comfortable. Family physicians in the city were and are extremely busy, and although some had offered to be available for immediate consultation if needed, this was only an *ad hoc* arrangement. Lack of medical coverage in the community was not felt by study investigators to be an issue. Subjects in

either group in the study were free to attend the ER or doctor's clinic as they thought necessary, so their choices were not constrained by participation in the study.

Study protocol required that referral to the QR project be made by ER physicians and consultant physicians only. In retrospect it might have been better, although harder to standardize, if family physicians had been part of the referral mechanism. Another possibility might have been to request participation in the study at the point patients were referred to Community Health; none of the patients who could have been enrolled at this stage would have fallen into the group which would have been admitted to hospital immediately, but as it was, fewer than 15% of those referred came into this category.

The passage of the Advance Health Care Directives Act on July 1st 1995 was a further setback for the QR study. This act forbids the participation in research studies of subjects whose comprehension is inadequate to give informed consent for participation, unless they have previously, while competent, indicated formally in writing their willingness to take part in such studies. Proxy consent by spouse, relatives or close friends is not sufficient to allow participation. As the Act now stands, it has been construed to apply to all forms of research whether invasive or not, unless the primary purpose is treatment rather than research. From the viewpoint of the study, it was unfortunate for two reasons: firstly, QR intervention had the potential to benefit incompetent persons and their families by providing short-term care and respite and encouraging discussion re the options for long-term arrangements. A number of patients already enrolled in the study fell into this category.

Secondly, because QR nurses had to explain to ER staff that from July 1st onward such patients, although meeting all other study criteria, could not be accepted into the study, an additional negative feeling toward the study was created in ER staff.

It is possible that sicker patients are more likely to refuse to take part in studies. Except for two individuals, available information on the management of patients who did not wish to take part in the QR study does not support this as the reason for their refusal. However, the patients had waited some hours in the ER before someone had decided that they were potential candidates for the QR study. It is possible that some of those people who did not wish to take part in the study may have been put off by the prospect of having to wait even longer in the ER.

In considering reasons for poor enrolment, it is clear that many factors were beyond the control of the investigators. Running the project at three sites was difficult. Dependence on extremely busy ER staff who were not part of the QR team to generate potential study subjects was a weakness of design which was not recognized in the feasibility study, because this was a chart review only. It is possible that a research nurse based in each ER could have brought eligible patients to the attention of ER staff, or perhaps better, that funds could have been provided for an additional ER nurse to be enlisted for this purpose.

Feedback from ER and Community Health staff indicates that a number of them felt that the project was superimposed on them when they already had more than enough pressure in their daily work. The QR team viewed the early conclusion of the project as a lost opportunity to give the QR concept formal evaluation.

Chapter 5

5.1 Care in the community: predictors of success: the previous chapters present the rationale, method and major results of the Quick Response study. This chapter is concerned with identification of factors which predict a successful intervention by Community Health.

The definition of success in the context of the community is elusive: what may be counted as a failure in one case may be success in another. Health care is superimposed on the natural history of diseases acute and chronic, and interacts with patient personality and social relationships. Intervention by health care providers does not always result in recovery of health. Given the wide variety of problems which lead patients to attend Emergency Departments, it is not surprising that predicted outcomes will range from full or partial recovery to marked incapacity or death. A successful Community Health intervention may enable a sick person to remain in his own home in spite of increasing disability, or may assist a patient more comfortably to the grave while lessening some of the burden on family caregivers. A positive outcome is only a partial success in terms of delivery of health care if the same end-result could have been achieved at a lower cost.

On the other hand, saving provincial dollars at the (often unmeasured) expense of the physical and psychological health of the caregiver is not only callous but in the long-run may be more costly: many caregivers are themselves elderly and perhaps fragile, and the extra burden may prove more than they are able to sustain.

Success was therefore defined in each case as achieving the best possible outcome, given the problem(s) documented in the Emergency Room referral to QR. Possible outcomes were defined in terms of need for future Community Health services:

(1) return to former function or better within the 2 week timeframe of QR;

(2) substantial improvement to the point of being appropriately cared for by regular Community Health services. Another possibility concerns those patients and their families who, because of illness or other crisis, have reached a point where long-term decisions about future care have to be made. A satisfactory outcome in such a case would be (3) the provision of services, such as home-making and professional support, and information about the options available, giving time to allow the patient and family to reach acceptable decisions about the long-term arrangements. Category (4) was provision of support for patients with terminal disease who wished to remain at home. A further category: return of patient or patient and family unit to independence (*i.e.* in terms of needing support from Community Health), was felt to overlap substantially with the first outcome listed, so the two were combined to read: "return to former function or better within the two-week timeframe of QR, and/or patient or patient and family unit unlikely to require additional services in the near future".

In order to identify an appropriate goal for each subject*, a physician who was neither involved with the patient's ER care nor aware of subsequent progress considered the information documented on the QR referral form. This included the presenting diagnosis and pertinent medical history, age, marital status, present occupation, household composition, and the referring physician's orders (services requested, and medications).

* One subject was enrolled twice, (once in the pilot phase and once in the randomization phase) with a different presenting diagnosis, and a different expectation identified each time. The major co-existing disease was the same. Unless stated otherwise, where analyses use subjects from both phases of the study, this subject is represented twice.

Using the same information, goals for each subject were independently determined by an experienced nurse. Neither the physician nor the nurse had the advantage of seeing the subjects when they attended the ER, and information recorded on the study referral forms did not always present a complete picture of the patient and his or her problems. Thus interpretation of the information provided sometimes led to a difference of opinion as to the appropriate goal. Agreement between the nurse and physician as to a realistic expectation in each case is shown in Table 5.1. Overall there was agreement in 60/82 cases (73%); this gives Cohen's κ, the agreement beyond that expected by chance, 0.54 (95% confidence interval 0.38 - 0.70). In the 22 cases where the nurse and physician had identified different goals, consensus decisions were reached after discussion.

Where possible, data reflecting subjects' health status at two and thirteen weeks after their ER visit were obtained from Community Health records. These records only cover time periods during which the agency is responsible for care, so in some cases, *e.g.* when patients were admitted to hospital, the required information had to be looked for elsewhere. Other instances where confirmation of progress was sought from different sources, such as hospitals or subjects' families, include three cases where records did not present a clear picture of status at two weeks, and five individuals assigned to SC who did not receive Community Health services.

Table 5.1: Comparison of Goals of Treatment identified by nurse and physician

RN → MD ↓	Group 1	Group 2	Group 3
Group 1	39	9	3
Group 2	2	16	5
Group 3		3	5

Group 1: return to former function or better within the 2 week timeframe of QR, and/or patient or patient and family unit unlikely to require additional services in the near future;

Group 2: substantial improvement to the point of being appropriately cared for by regular Community Health services.

Group 3: the provision of services, such as home-making and professional support, and information about the options available, giving time to allow the patient and family to reach acceptable decisions about the long-term arrangements to be made.

The Group 4 objective: provision of support for patients with terminal disease who wished to remain at home, was not chosen as appropriate for any subject.

Achievement of goals at two weeks was classified as achieved, partly achieved, or not achieved. For each patient, the set goal was rephrased as a question, and only if the answer was an unequivocal "yes" was it judged that the goal had been achieved. For example "did the patient return to his former level of function, and/or does the patient or patient and family unit require no Community Health services after two weeks (or no more than previously needed)?" Failure to achieve the goal included admission to hospital within the first 14 days after the ER visit (where admission was not part of Standard Care); or deterioration in clinical status, or no progress towards the identified goal. Partial success was judged to have been attained at two weeks if it was documented clearly that considerable progress towards the goal had been made. Partial success was defined so that it was much closer to success than failure on the continuum of achievement. As a measure of reliability, a lay person, not connected with the QR project, also made these judgements, independently, based on summaries (without names or identifying details) of the relevant Community Health records and in some cases, as noted above, using other reliable information.

Table 5.2 shows how the two assessors rated the two-week outcome in relation to the goal identified for each subject. The measure of agreement was 88%, with κ 0.80, (95% confidence interval 0.69 - 0.91). Disagreements were resolved by referring to the additional information provided by the Barthel and other questionnaires.

Table 5.2. Accord in rating achievement of goals at two weeks

GB → LP ↓	Successful	Partly successful	Not successful
Successful	41	7	0
Partly successful	0	11	3
Not successful	0	0	20

Another method of establishing whether the identified goal had been met was to use Barthel scores. Scores reflecting prior function (before the episode which provoked the ER visit), function at the time of the ER visit, and 2 weeks after were compared, to provide another measure of progress. Independence in activities of daily living (ADL) gives some indication of health, particularly in the elderly or incapacitated.

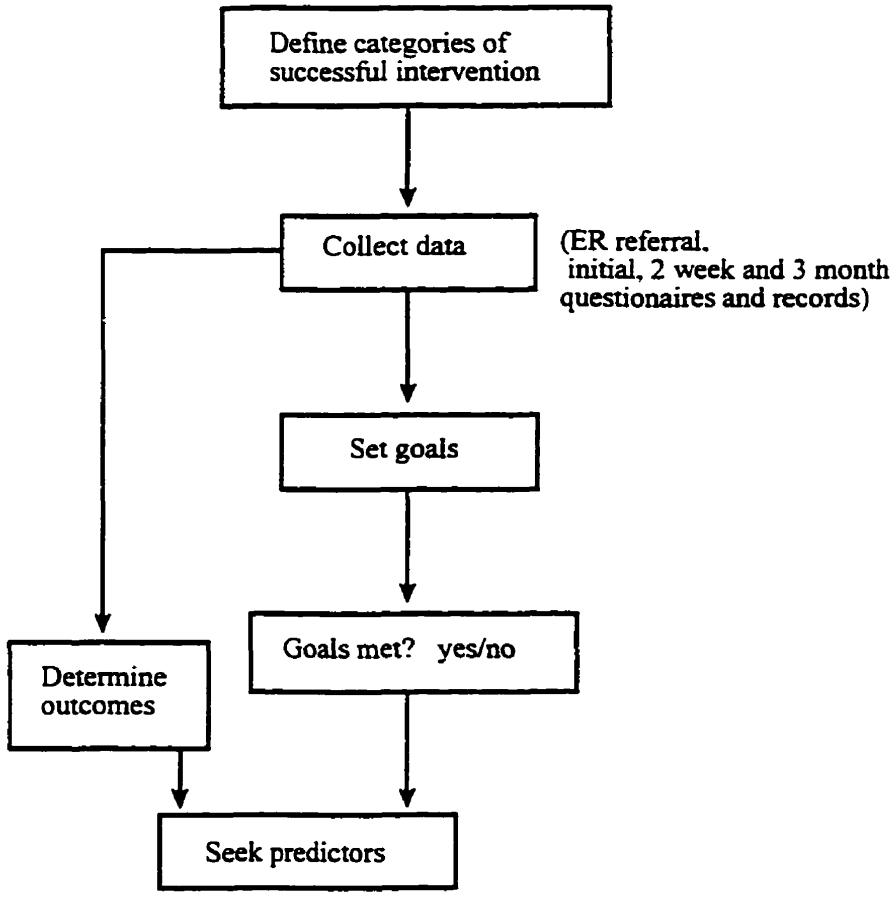
To confirm "return to prior function", the Barthel score at two weeks had to equal or surpass the score reflecting function before the ER visit; otherwise, change was not accepted as clinically significant unless the Barthel score differed by at least 7% from that obtained in the ER. In 21 cases (26% of all subjects) Barthel scores in the ER were ≥ 99%, and were therefore not useful in judging improvement. For these subjects, additional information was sought from Sickness Impact Profile (SIP) scores and the subject's self-rating of health in the SF-36.

Questionnaire data was compared with Community Health records, to see if information about each subject from the different sources was compatible. In a few cases, where information recorded by Community Health staff referred to a different time point (e.g. three or four weeks rather than two weeks after the ER visit) greater weight was given to questionnaire data, and judgement of how well a subject achieved the set goal was adjusted to reflect the additional information. Figure 5.1 outlines the process of goal determination, and judgement of goal attainment by each subject.

Characteristics of subjects in each group (achieved, partially achieved and failed to achieve the expected goal) were examined to identify predictors of success or failure. The chi-squared test was used to evaluate the association of success or failure with categorical variables such as sex and marital status. The Wilcoxon Rank Sum test was used to examine the predictive value of the ER Barthel score. Logistic regression was used to investigate how well information available in the ER such as age, diagnosis and Barthel score, can predict success in meeting goals.

A premise here is that a valid identification of goals can be made using the information provided on the ER referral forms, without seeing the patients. Recognizing the difficulties inherent in such prediction (and accepting that identified goals in some cases may have been unrealistic or inappropriate), logistic regression was also carried out to analyse how well health outcomes can be predicted by these factors.

Figure 5.1. Protocol for determining predictors of goal attainment and health outcome



Chapter 6

6.1 Achievement of goals: With a few exceptions (8/82), the picture of each subject given by the various questionnaires was in agreement with that documented by Community Health records. Discrepancies were in most cases minor, several relating to the timing of recovery of independence in daily function. In one case, the accuracy of self-reported ability prior to the ER visit was in doubt. Another case, while successful overall, was only a partial success in economic terms, because the solution arrived at was very costly. Four patients declined to answer study questionnaires at two weeks, although they continued to receive Community Health services as planned. They consented to the use of their health records and MCP billing data, so their status was known. Table 6.1 summarizes the goals identified, and the degree to which these goals were achieved two weeks after the ER visit.

Table 6.1 Achievement of goals at two weeks

Goal identified by MD and RN	n	Achieved	Partly achieved	Did not achieve
1. Return to former function or better, or return of patient/family unit to independence	47	22 (46.8%)	15 (31.9%)	10 (21.3%)
2. Substantial improvement: can be cared for by routine Community Health services	24	12 (50.0%)	1 (4.2%)	11 (45.8%)
3. Provision of time to allow decision-making re ongoing care	11	8 (72.7%)	1 (9.1%)	2 (18.2%)
TOTAL	82	42	17	23

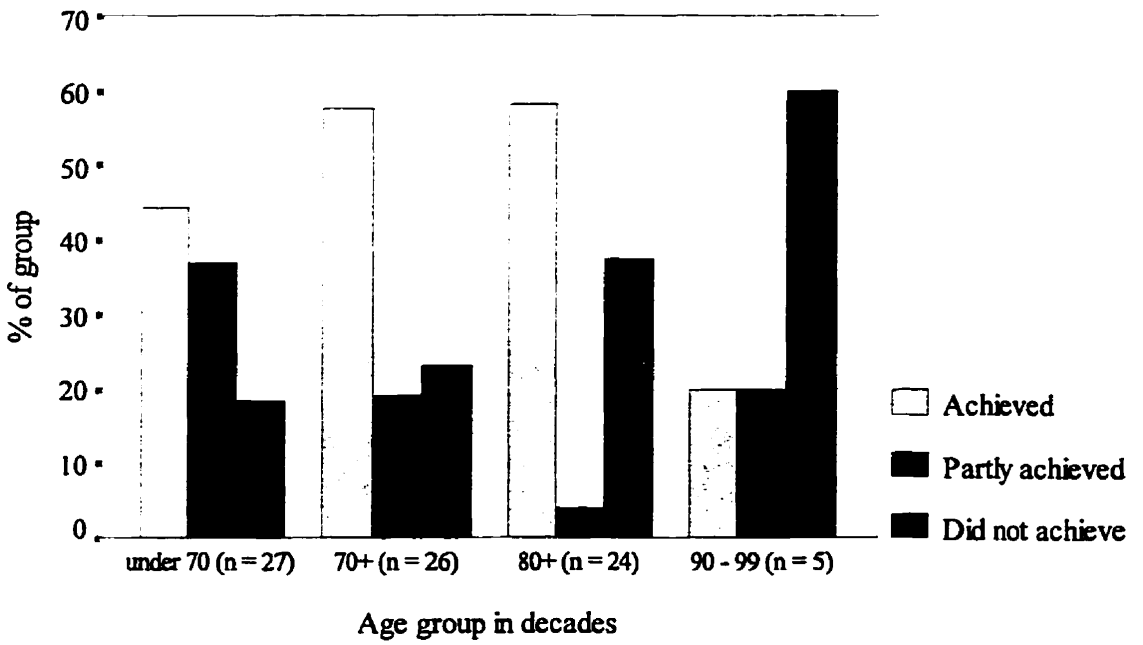


Figure 6.1 Age group and achievement of goals at two weeks

Figure 6.1 shows how well subjects in the various age groups fared in relation to the goals set for them. Subjects aged 80 years or older were twice as likely to be unsuccessful at two weeks (41%) as those aged under 80 years (21% of whom had not reached their goal at two weeks). The failure rate in the five subjects aged 90 years or over was even greater (60%). The median age of subjects who achieved the set goal was 76 years; of those who had made some progress, 54 years; and of those who did not achieve the set goal, 80 years. With the exception of the small group of 8 patients aged 60 - 69 years, there is an almost linear relationship between age and percentage who were unsuccessful at two weeks.

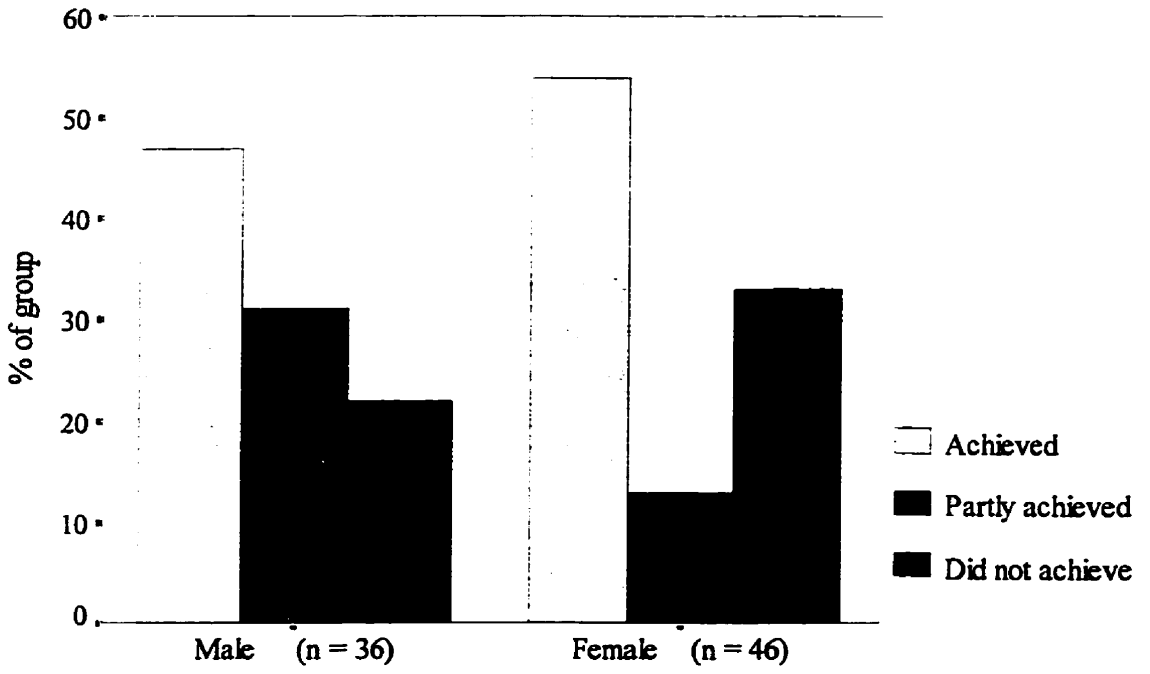


Figure 6.2 Sex and achievement of goals at two weeks

Figure 6.2 compares the achievement of male and female study subjects: differences were not statistically significant (chi-squared test, $p = 0.14$).

Marital status may influence achievement of goals at two weeks, but the differences shown in Table 6.2 are not significant, (chi-squared $p = 0.27$). Support from family and friends was a clearer determinant of whether subjects would achieve the goals set for them, see Figure 6.3. Support from family and friends was judged positive when the QR nurse in the ER documented that there appeared to be a good relationship between a subject and the accompanying family member, or that regular assistance with ADL was provided by family members or friends. A lack of support was identified in the following situations: when an abusive relationship was documented; when the person with whom the subject lived was unable to help with ADL, but no-one else was named as a contact; when the contact named on the QR referral lived off the Avalon Peninsula; and when assistance with ADL was provided only by paid homemakers. Subjects without support were not accompanied to the ER, and most (9/12) went to and from the ER by ambulance or taxi.

84% of subjects could rely on encouragement and help from family or friends. These people were five times more likely to meet the set goals than subjects who lacked support, chi-squared $p < .005$ (with Yate's correction for expected number < 5 in two cells).

Table 6.2 Marital status and achievement of goals at two weeks

	n	married n = 30	widowed n = 29	single n = 18	divorced/separated n = 5
Achieved	42	14 (47%)	16 (55%)	10 (56%)	2 (40%)
Partly achieved	17	10 (33%)	2 (7%)	4 (22%)	1 (20%)
Did not achieve	23	6 (20%)	11 (38%)	4 (22%)	2 (40%)

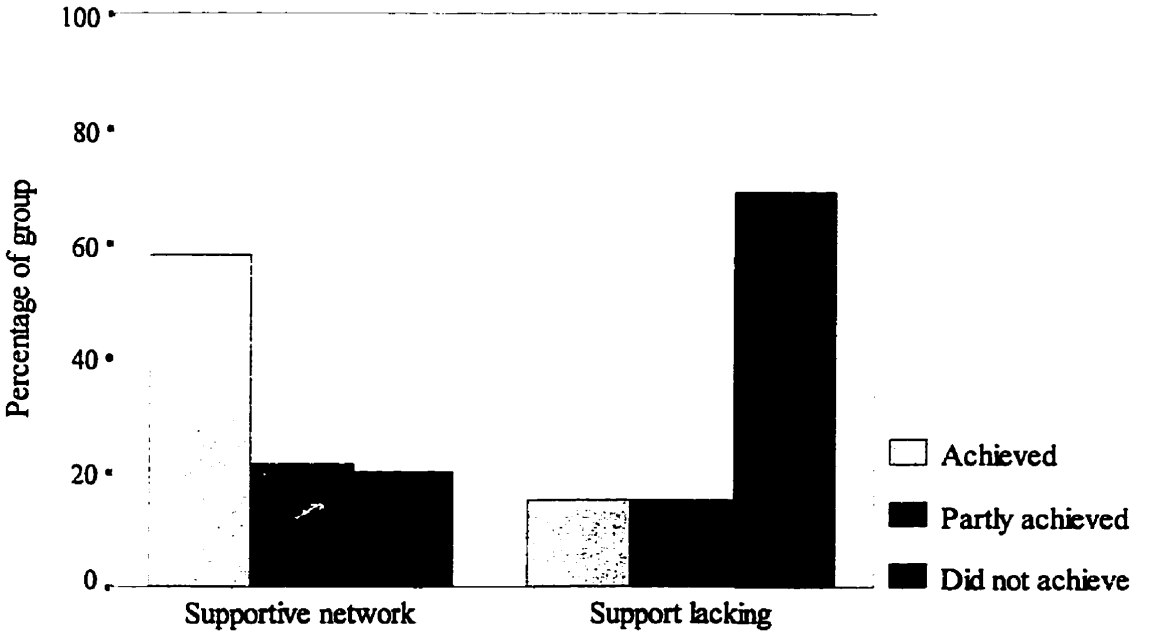


Figure 6.3 Support from family and friends and success in meeting goals at 2 weeks

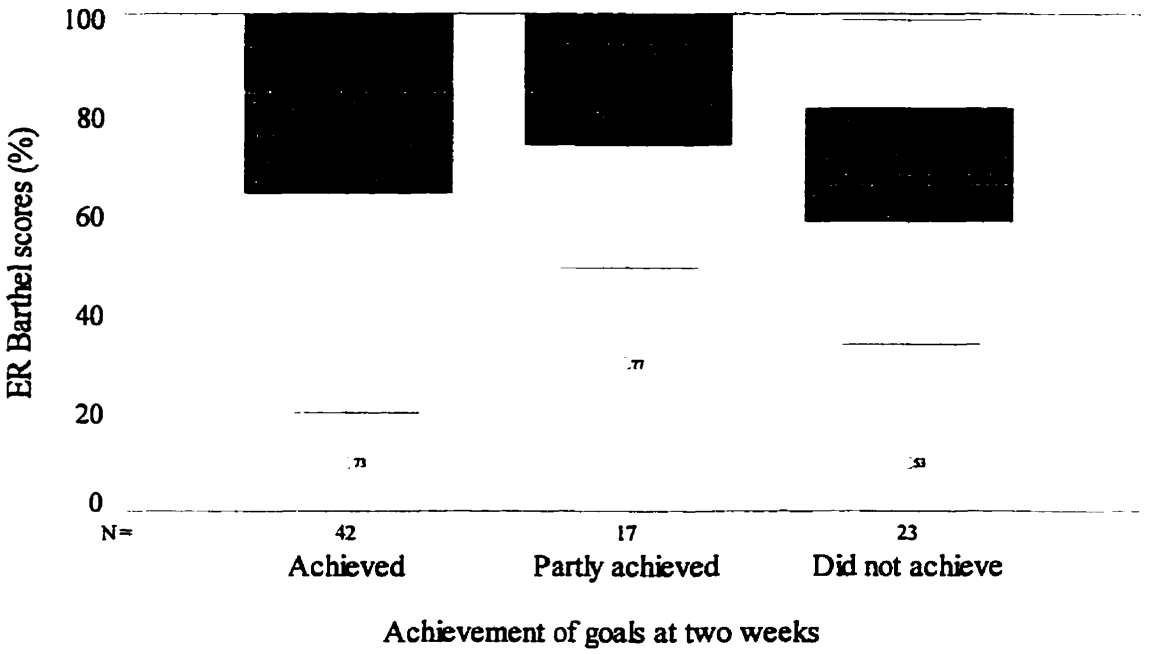


Figure 6.4 Initial (ER) Barthel scores and achievement of goals: box and whisker plot showing median, 25th and 75th percentile scores

There was a clinically and statistically significant difference in ER Barthel scores between the 23 subjects who did not succeed and the 59 who either succeeded (42) or made some progress (17) in the first two weeks after their ER visit ($p = .01$, Wilcoxon 2-tailed test), Figure 6.4.

Taking 70% as a cut-off, the 54 subjects with ER Barthel scores $\geq$ 70% were 1.5 times as likely to achieve, or to make progress toward achieving, the set goal at two weeks as those who scored less than 70%. Conversely, the 28 subjects who had a Barthel score less than 70% were 2.5 times more likely not to achieve their expected goal (chi-squared $.05 > p > .01$). Although the Barthel score would appear to have some predictive value, once other factors such as age and diagnoses were factored in, the ER Barthel score did not discriminate further between those who would or would not succeed at two weeks.

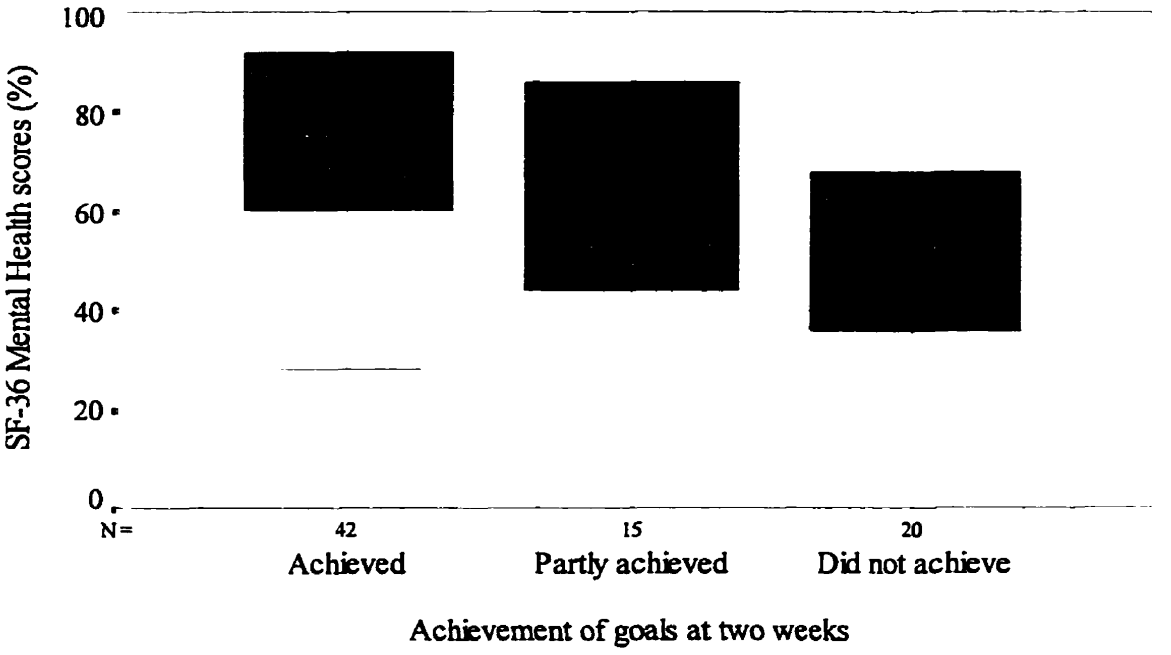


Figure 6.5 Initial SF-36 Mental Health scores and achievement of goals: box and whisker plot showing median, 25th and 75th percentiles

Comparing those who failed to achieve the two-week goal with those who either achieved or made some progress toward achieving their goal, there was a significant statistical difference in the initial SF-36 MH score (Wilcoxon rank sum test, $p = .002$), Figure 6.5. Although the difference was not clinically significant according to the literature criterion which distinguished those with and without psychiatric disease, it does serve here to distinguish between those who achieved or made some progress towards their goal at two weeks.

5 subjects (2 who partly achieved and 3 who failed to meet the goal) did not answer this questionnaire (the first 2 refused, the other 3 were too sick).

Table 6.3 Presenting diagnosis and achievement of goals at two weeks

	n	Ortho n = 40	Infection n = 13	Urinary n = 5	Cardio-resp n = 11	Other n = 13
Achieved	42	16 (40%)	8 (61.5%)	4 (80%)	7 (64%)	7 (54%)
Partly achieved	17	11 (27.5%)	4 (31%)			2 (15%)
Did not achieve	23	13 (32.5%)	1 (8%)	1 (20%)	4 (36%)	4 (31%)

"Ortho" includes fractures, neck, back and other soft tissue injuries and degenerative joint disease. "Infection" includes skin and tissue infection. "Urinary" includes urinary retention and urinary tract infection. "Cardio-Resp" indicates acute respiratory problems in most cases superimposed on chronic respiratory conditions, or cardiac insufficiency.

"Other" includes neurological problems, weight loss, dehydration, weakness and anxiety.

Table 6.3 gives a summary of the diagnoses documented as the reason for the ER visit, and the degree of success for each group in achieving goals at two weeks. Numbers in each group are too small too draw any conclusion, although it appears that subjects who came to the ER with infections generally responded well to antibiotic therapy. Urinary tract infections and urinary retention were indicative of underlying problems, commonly prostatic hypertrophy, but in most cases the short-term problem responded to treatment.

Classification of diagnoses as in Table 6.3 did not help in predicting which subjects would achieve their goals at two weeks, but the presence of one or more adverse diagnoses (noted on the ER referral form, and listed below) was clearly associated with lack of success: (they were 4.7 times more likely to fail than subjects without one of these diagnoses, chi-squared $p < 0.001$). Adverse diagnoses were identified as congestive heart failure or left ventricular failure; poorly controlled hypertension; chronic obstructive pulmonary disease; peripheral vascular disease; acute cerebrovascular accident; old cerebrovascular accident or other neurological diagnosis with consequent impairment of mobility; dementia. Twenty-seven subjects were identified in the ER as having one or more of these diagnoses, although in most (20/27) cases this was not the presenting problem. Not included in this list of adverse diagnoses are stable ischemic heart disease, controlled hypertension, diabetes mellitus whether insulin dependent or not, carcinoma, arthritis or degenerative disc disease, psychiatric disease or spousal abuse.

Twenty-eight subjects came to the ER having fallen; eventually 18 of them recovered full function although within the two-week timeframe of QR only 7 had recovered fully. A fall may suggest underlying problems: on review, falls were classified as uncomplicated (no serious concurrent diagnoses) or complicated (accompanied by one or more of the problems noted above). The 19 patients whose fracture or soft tissue injury was uncomplicated were seven times more likely to achieve or partially achieve the two-week goal set for them than the 9 whose fall was

complicated by serious medical problems, (Yates' corrected chi-squared $p < 0.005$).

There were 9 subjects with a diagnosis of dementia; they were much less likely to achieve the goals set than other subjects: (29% reached their goal by comparison with 55%) and conversely much more likely not to reach their goal (56%by comparison with 22%).

Young subjects (aged less than 50 years) who presented with acute back pain eventually made a full recovery although at two weeks only two of the five had fully achieved their goal; the other three, although they had made some progress, were still moving very cautiously. Nine older subjects (aged more than 50 years) also presented with back pain, but they all had accompanying medical problems of varying severity: nearly half (44%) had compression fractures indicative of osteoporosis. At two weeks, four of the nine (44%) had achieved their goal and three more had made some progress, but at three months, only three (33%) of these older subjects had made a full recovery, while three had improved but still had problems, two were clearly worse and one was dead.

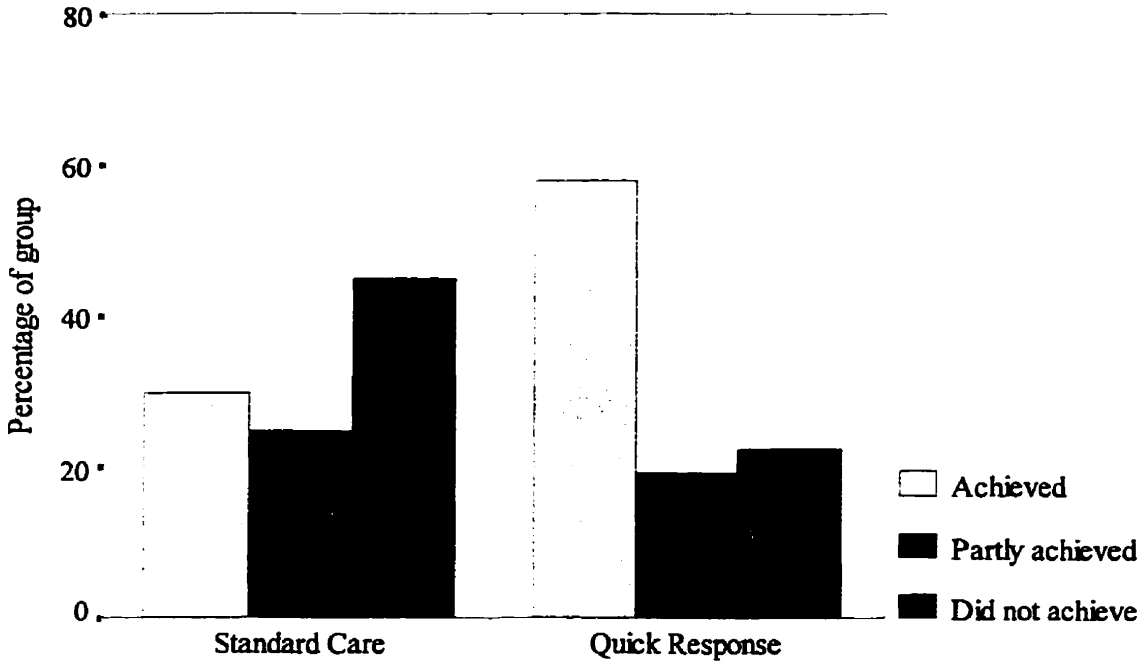


Figure 6.6 Type of care and success in meeting goals at two weeks

Figure 6.6 shows the association between the type of care received (QR or SC) and achievement at two weeks. Those who received QR were somewhat more likely to achieve or make some progress toward achieving their goals than those who received SC (chi-squared, $0.05 < p < 0.1$, with Yate's correction for the low expected number in one cell). This was not a strong relationship and once other variables such as age and the presence of an unfavourable diagnosis were factored in, type of care was not identified in the regression analysis as a predictor of two-week success (see page 107).

6.2 Health outcomes: the two-week goal selected for each subject reflected expert judgement of the likely outcome, based on the facts available in the ER, as documented on the QR referral form. Criteria for achievement were defined as rigorously as possible (page 82), nevertheless the decision as to degree of achievement was in ~ 15% of cases a matter of opinion. By contrast, health status in the majority of cases is on record and can be recognized more objectively.

Tables 6.4 and 6.5 show health outcomes at two weeks and three months, by study group: there was no significant difference in outcome between SC and QR subjects. Table 6.6 shows the status of study subjects at three months, grouped according to achievement of goals at two weeks. This shows that if at two weeks subjects had achieved the goals set for them or had made some progress toward these goals, they were more than twice as likely to be fully recovered by three months as those who had made no progress toward attaining their two-week goal. Conversely, those who had made no progress were more than 3½ times more likely to have a poor outcome (chi-squared, $p < .005$).

Figures 6.7 - 6.9 show how two-week health outcome varied with age group, sex, and support from family and friends; Tables 6.7 - 6.9 show outcome at two weeks according to marital status, and outcome at two weeks and three months by presenting diagnosis. These are the same variables discussed with respect to goal achievement, and in general they exert their effects in the expected directions, *e.g.* as a group, older people tended to fare less well (Figure 6.7); the 32 subjects who at three months were fully recovered had a mean age 10 years lower than other subjects (median difference 3½ years). No significant difference in outcome was seen between males and females (Figure 6.8). Figures 6.10 and 6.11 show ER Barthel scores according to outcome at two weeks and three months.

Table 6.4 Study group and health outcomes of study subjects at two weeks

Outcomes	Non-randomized	Randomization Phase	
	Quick Response n = 40	**Standard Care** n = 20	**Quick Response** n = 22
Essentially recovered	16 (40.0%)	2 (10.0%)	5 (22.7%)
Some improvement	8 (20.0%)	10 (50.0%)	11 (50.0%)
No change	8 (20.0%)	2 (10.0%)	2 (9.1%)
Worse	7 (17.5%)	6 (30.0%)	4 (18.2%)
Death	1 (2.5%)	0	0

Categories here are described with reference to the day of the ER visit, thus all categories except "essentially recovered" reflect deterioration from health status prior to the ER visit.

Table 6.5 Study group and health outcomes of study subjects at three months

Outcomes	Non-randomized	Randomization Phase	
	Quick Response n = 40	**Standard Care** n = 20	**Quick Response** n = 22
Recovered	16	8	8
Some improvement	5	3	2
Immediate problem solved:			
underlying problem remains	7	3	5
or **new problem presents**	3		
No change	1	2	3
Worse	4	4	3
Moribund	1	0	1
Death during study	3	0	0

Table 6.6 Status of subjects at three months, grouped by achievement of goals at two weeks

2/52 goals	Recovered	Improved + Immediate problem solved, but underlying problem persists, or new problem presents	No change	Worse	Dead	Where
Achieved n = 42	22 (52%)	4 (10%) + 10 (24%)	3 (7%)	3 (7%)		At home: 38 Palliative Care: 1 Long-term care: 2 Lost to follow-up: 1
Partly achieved n = 17	9 (53%)	5 (29%)		3 (18%)		At home: 15 (of whom 1 awaiting hospital admission) In hospital: 2
Did not achieve n = 23	5 (22%)	1 (4%) + 4 (17%)	3 (13%)	7 (30%)	3 (13%)	At home: 14 (of whom 2 applying for long-term care) In hospital: 3 Long-term care: 3

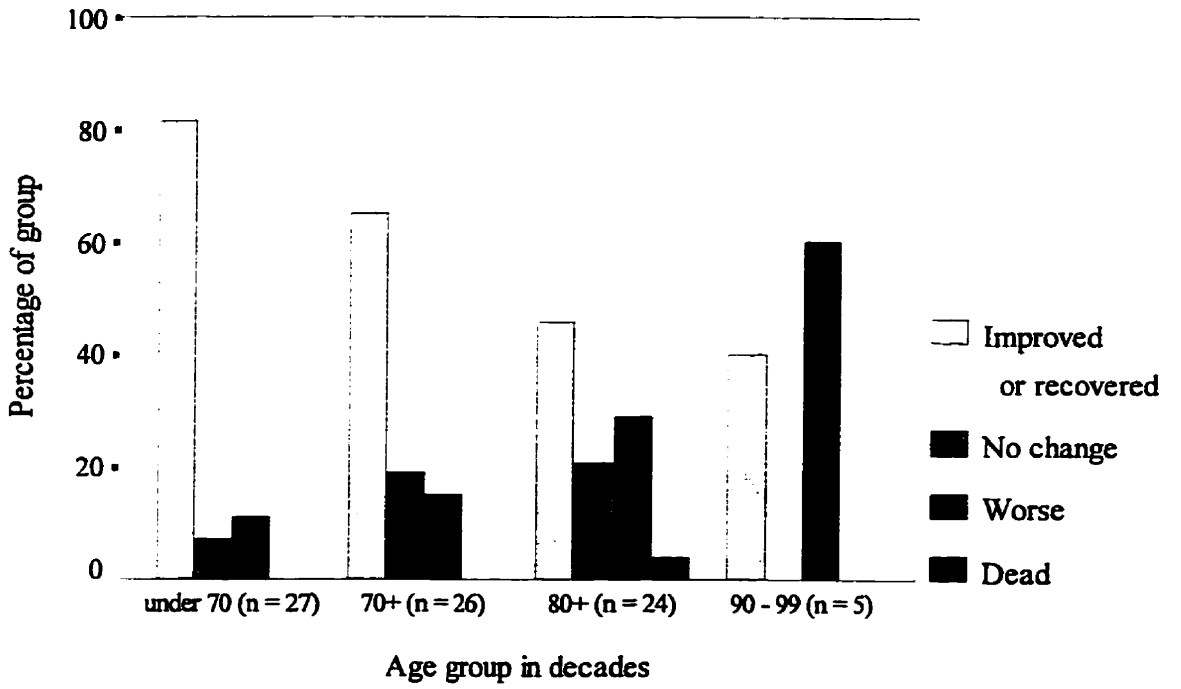


Figure 6.7 Age group and outcome at two weeks

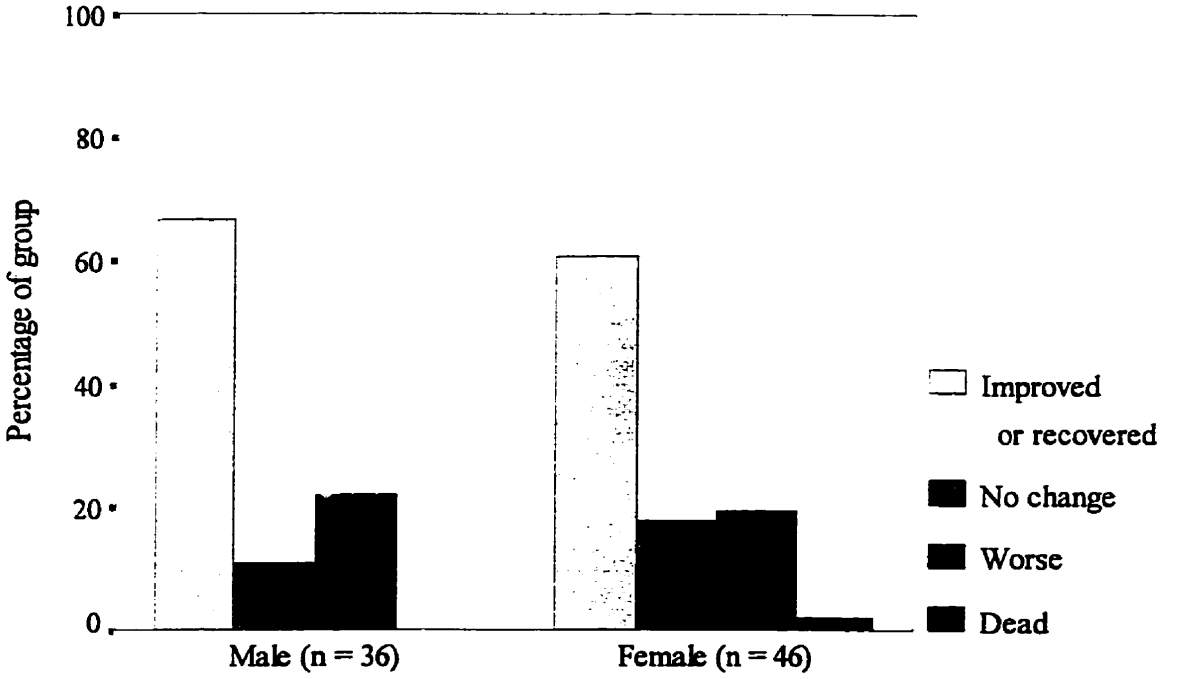


Figure 6.8 Sex and outcome at two weeks

Table 6.7 Marital status and outcome at two weeks

	n	married	widowed	single	divorced/separated
		30	29	18	5
Improving or recovered	52	19 (63%)	17 (59%)	13 (72%)	3 (60%)
No change	12	3 (10%)	7 (24%)	2 (11%)	
Worse	17	8 (27%)	4 (14%)	3 (17%)	2 (40%)
Dead	1		1 (3%)		

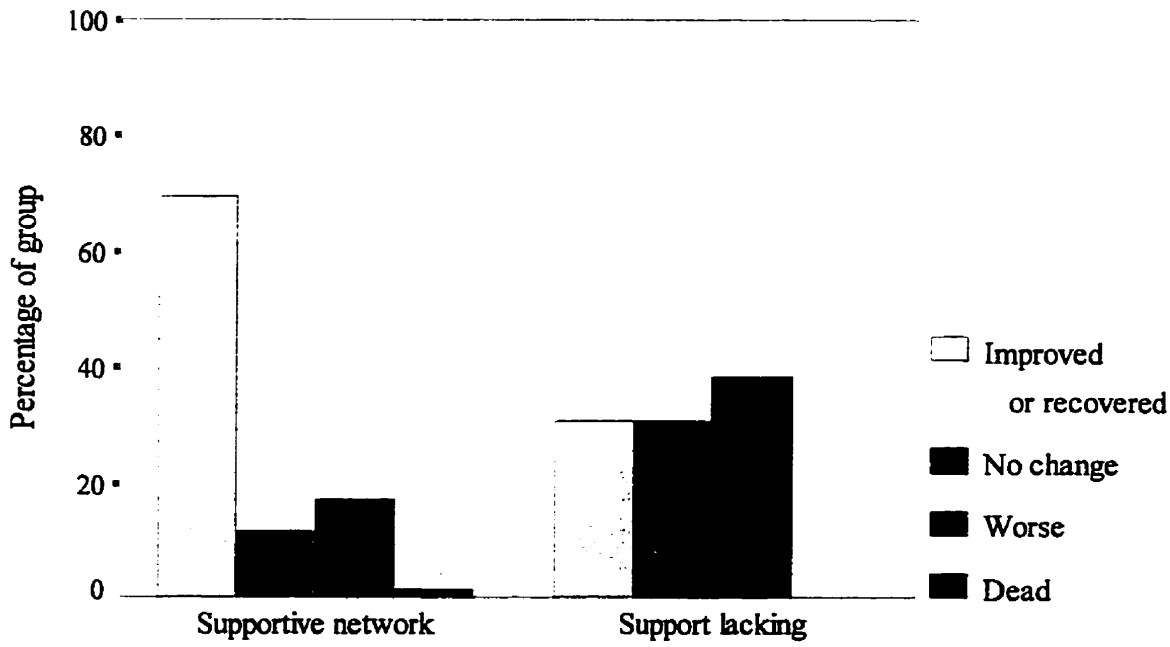


Figure 6.9 Support from family and friends and outcome at two weeks

Table 6.8 Presenting diagnosis and outcome at two weeks

	n	Ortho n = 40	Infection n = 13	Urinary n = 5	Cardio-resp n = 11	Other n = 13
Recovered or improved	23 29	9 (22.5%) 18 (45.0%)	6 (46.2%) 5 (38.5%)	1 (20.0%)	4 (36.4%) 1 (9.1%)	4 (30.8%) 4 (30.8%)
No change	12	5 (12.5%)	1 (7.7%)		3 (27.3%)	3 (23.1%)
Worse **Dead**	17 1	7 (17.5%) 1 2.5%	1 (7.7%)	4 (80.0%)	3 (27.3%)	2 (15.4%)

"Ortho" includes fractures, neck, back and other soft tissue injuries and degenerative joint disease. "Infection" includes skin and tissue infection. "Urinary" includes urinary retention and urinary tract infection. "Cardio-Resp" indicates acute respiratory problems in most cases superimposed on chronic respiratory conditions, or cardiac insufficiency. "Other" includes neurological problems, weight loss, dehydration, weakness and anxiety.

More than half the subjects who came to the ER with orthopedic problems either recovered or made good progress in the first two weeks. Of the 40 patients whose presenting diagnosis was orthopedic, 10 had one or more of the unfavourable diagnoses listed on page 94; these subjects were over five times as likely to deteriorate in the first two weeks as those whose general health was good or whose medical problems were well controlled or not active, chi-squared $p < .005$, and five times as likely to be worse or dead at three months, $p < .05$ (Table 6.9).

Table 6.9 Presenting diagnosis and outcome at three months

	n	Ortho n = 40	Infection n = 13	Urinary n = 5	Cardio-resp n = 11	Other n = 13
Recovered	32	18 (45.0%)	5 (38.5%)		6 (54.6%)	3 (23.1%)
Improved	10	7 (17.5%)		1 (20.0%)	1 (9.1%)	1 (7.7%)
Immediate problem solved: underlying problem remains or new problem presents	18	7 (17.5%)	7 (53.9%)	2 (40.0%)		2 (15.4%)
No change	6		1 (7.7%)		1 (9.1%)	4 (30.8%)
Worse	13	6 (15.0%)		2 (40.0%)	2 (18.2%)	3 (23.1%)
Dead	3	2 (5.0%)			1 (9.1%)	

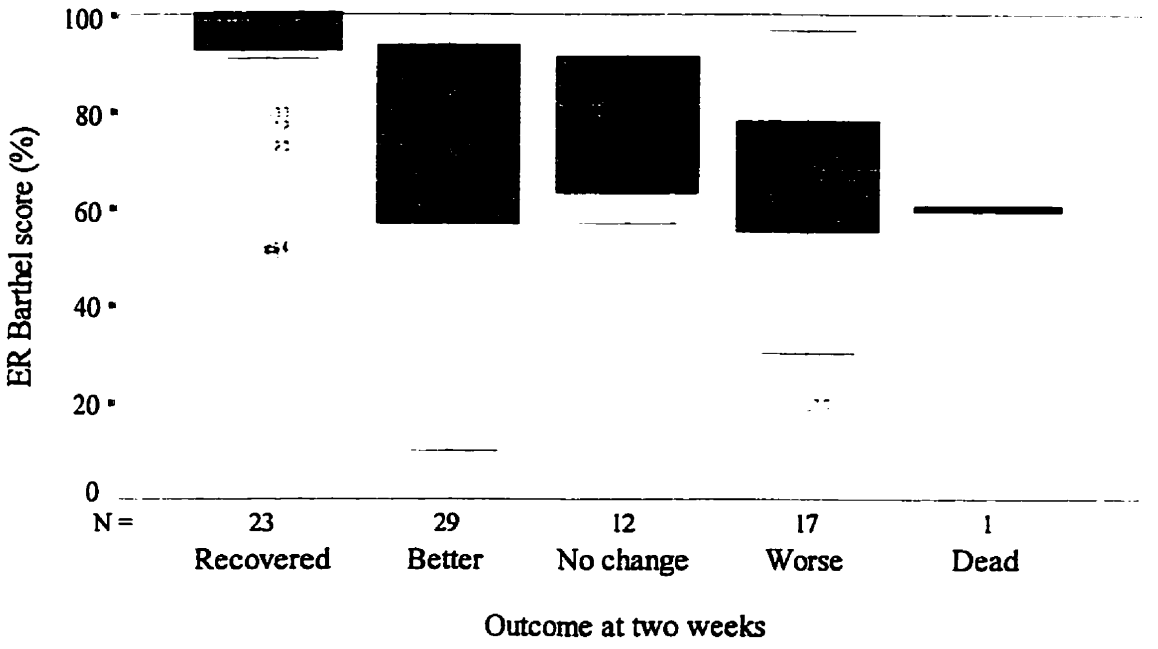


Figure 6.10 Initial (ER) Barthel scores and outcome at two weeks: box and whisker plot showing median, 25th and 75th percentiles

There was a wide difference in ADL ability as documented in the ER between those who would deteriorate over the next two weeks and those who either improved or whose health status was unchanged, $p = .001$.

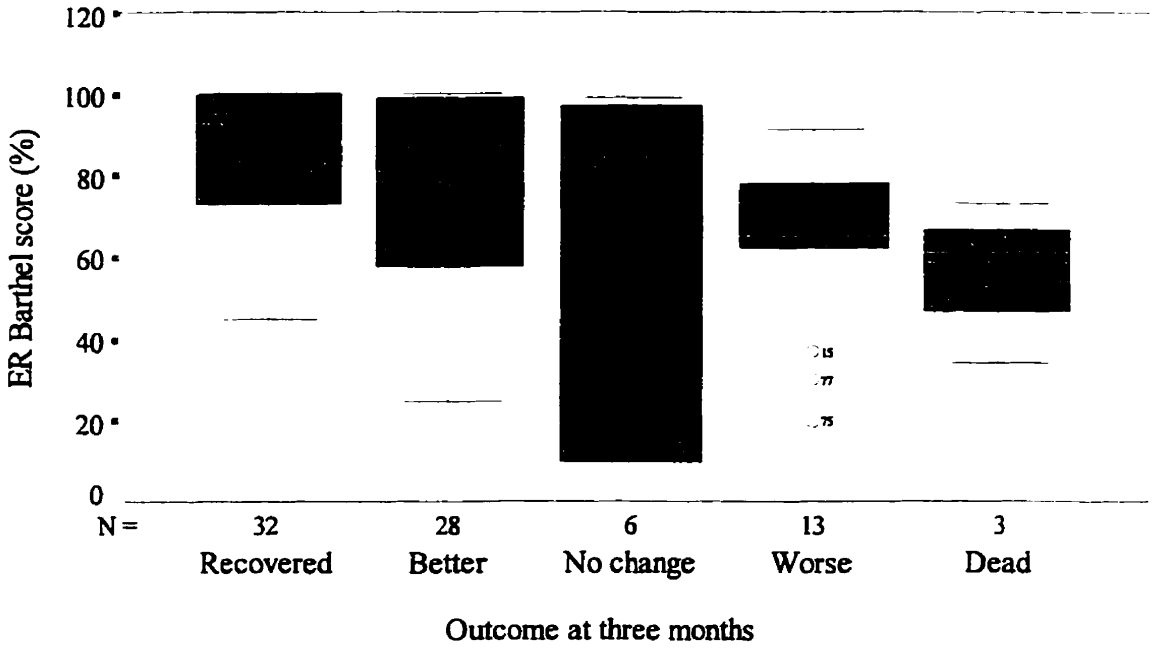


Figure 6.11 Initial (ER) Barthel scores and outcome at three months: box and whisker plot showing median, 25th and 75th percentiles

Figure 6.11 shows the wide difference in ER Barthel scores between the 12 subjects who would be fully recovered by three months, score 84.9 ± 17.8, median 94, and the 12 who at three months were either much worse (9) or dead (3), score 56.7 ± 21.3, median 62, Wilcoxon rank sum test p = .004. Median ER Barthel scores for those whose health at three months would be essentially the same as at their ER visit, somewat recovered or rather worse, ranged from 79 - 83; means ranged from 63.7 (unchanged at three months) to 78.6 (acute problem resolved but underlying condition unchanged, or recovered but now has a new problem).

6.3 Logistic Regression*. Variables which appeared to have some bearing on the achievement of goals at two weeks were analysed by logistic regression. These variables were age or age group; presence or absence of a supportive network; ER Barthel score; initial SF-36 Mental Health score; presenting diagnosis; presence or absence of one or more unfavourable diagnoses (listed on page 94); and study group (SC or QR).

Forward stepwise logistic regression identified three factors predictive of failure to achieve or make any progress towards achieving the two-week goal: lack of a supportive network, diagnosis (presence, not necessarily as the presenting problem, of one or more unfavourable diagnoses), and age. Neither the Barthel score recorded in the ER, nor the type of care given (SC or QR) contributed further towards identifying subjects who would fail to meet the two-week goals set for them. If adverse diagnosis was not factored into the analysis, a low score on the SF-36 Mental Health (MH) domain was also predictive of failure. In association with age and lack of support, a low score for this index, reflecting fragile mental health, (specifically general affect, behavioral-emotional control, anxiety and depression) appears to identify many of those subjects who had one or more unfavourable diagnoses. (An alternative if simplistic explanation might be that coming to an acceptance of poor health is stressful.) SF-36 MH data were missing for five subjects (two in the "partially achieved" group, who decided not to answer the questionnaires, and three in the "did not achieve" group, who were too sick to answer); inclusion of this variable in the regression analysis would have deleted these subjects from the model. It was therefore decided that the model would be more representative if the adverse diagnosis variable were used rather than the initial MH variable.

* See appendix pages 129 - 133 for a summary of calculations.

The regression equation predicts those who would not achieve the two-week goal:

$$\text{Probability of failure} = \frac{1}{1 + e^{-z}}$$

$$\text{where } Z = B_0 + B_1(\text{support}) + B_2(\text{diagnosis}) + B_3(\text{age in years})$$

The categorical variables supportive network and adverse medical diagnosis are coded 0.5 when present, and - 0.5 when absent,

e.g. if there is a supportive network of family and friends,

B_1(support) in the equation above is entered as 0.5 multiplied by - 4.1546 = - 2.0773.

Variables in the equation:

	B	df	Sig	R	Exp(B)
Support	-4.1546	1	.002	-.3498	.0157
Age (years)	.0857	1	.0337	.1605	1.0894
Diagnosis	2.9489	1	.0004	.3277	19.0846
Constant	-6.1259	1	.0415		

This model accurately predicts the achievement or lack of achievement of two-week goals for 66/82 (80%) of study subjects, taking the cutpoint of predicted probability at 0.5, such that < 0.5 predicts achievement and > 0.5 predicts failure.

Table 6.10 shows that the **sensitivity** of the model: correct prediction of those who would succeed divided by the number of all those who actually did succeed (true positive / true positive + false negative) was better at 83.1% than its **specificity**: correct prediction of those who would fail divided by all those who actually failed (true negative / true negative + false positive), 73.9%.

(This refers to the regression model with cutpoint of predicted probability at 0.5).

Table 6.10 Success in meeting goals at two weeks: predicted and observed

		Observed	
		Achieved or partly achieved	Did not achieve
Predicted	Achieve or partly achieve	49	6
	Will not achieve	10	17
		83.05% (sensitivity)	73.91% (specificity)

Correct overall: 80.49%

Although goal achievement and outcome at two weeks embody different concepts, there is clearly an association between them: there was 79% concordance between subjects who achieved or partially achieved their goals and those whose outcome was positive at two weeks, with $\kappa = 0.58$ (95% confidence interval 0.40 - 0.76).

Table 6.11 Outcome and achievement of goals at two weeks

Goal	Outcome		
	Improved + recovered	No change	Worse + dead
Achieved n = 42 (%)	13 + 21 (81.0%)	3 (7.1%)	5 + 0 (11.9%)
Partly achieved n = 17	12 + 2 (82.4%)	2 (11.8%)	1 + 0 (5.9%)
Not achieved n = 23	4 + 0 (17.4%)	7 (30.4%)	11 + 1 (52.2%)

Factors which predict lack of success in achieving goals at two weeks also predict poor outcomes at two weeks and three months, but the presence or absence of an unfavourable diagnosis is more important, and the ability of a supportive network to moderate the dependent variable, outcome, is less important than in the case of goal achievement.

The regression equation is configured to predict a positive outcome at two weeks:

$$\text{Probability of positive outcome} = \frac{1}{1 + e^{-z}}$$

$$\text{where } Z = B_0 + B_1(\text{support}) + B_2(\text{diagnosis}) + B_3(\text{age in years})$$

Variables in the equation: (with diagnosis and support coded 0.5 if present, - 0.5 absent)

	B	df	Sig	R	Exp(B)
Diagnosis	-3.1997	1	.000	-.4018	.0408
Age (years)	.0822	1	.0327	-.1594	.9211
Support	2.7684	1	.0038	.2516	15.9334
Constant	6.0093	1	.0364		

In terms of goodness of fit, the model for outcome at two weeks fits the data more comfortably than the model predicting achievement of goals at two weeks, with 68.600 by comparison with 51.267. The -2LL was 57.365, compared with 54.688 for goal achievement at two weeks.

Table 6.12 Outcome at two weeks: predicted and observed

		Observed: Positive	Observed: Negative
Predicted	Positive	48	6
	Negative	9	19
		84.2% (sensitivity)	76% (specificity)

Correct overall: 81.7%

In 12 cases where the two-week status was essentially unchanged from the ER status, the decision to classify the outcome as negative (6 subjects) or positive was made with reference to subsequent progress.

The regression equation for three months predicts positive outcome:

$$\text{Probability of positive outcome} = \frac{1}{1 + e^{-z}}$$

$$\text{where } Z = B_0 + B_1(\text{support}) + B_2(\text{diagnosis}) + B_3(\text{age in years})$$

Variables in the equation: (with diagnosis and support coded 0.5 if present, - 0.5 absent)

	B	df	Sig	R	Exp(B)
Diagnosis	-2.4048	1	.0006	-.3250	.0903
Age (years)	.0968	1	.0143	-.2071	.9078
Support	2.6028	1	.0060	.2437	13.5015
Constant	7.5346	1	.0109		

-2 LL 61.217 Goodness of Fit 60.361

Table 6.13 Outcome at three months: predicted and observed

		Observed	
		Positive	Negative
Predicted	Positive	52	8
	Negative	9	13
		85.3% (sensitivity)	61.9% (specificity)

Correct overall 79.3%

One case whose status at three months differed little from that at the ER visit was assigned to "outcome positive" because surgery to address the problem which provoked the ER visit had been carried out successfully, and although progress post-operatively was slow, the prognosis was much improved.

Logistic regression suggests that age exerts a greater influence on outcome at three months than at two weeks; the presence or absence of an adverse diagnosis is a little less important at three months than two weeks. A supportive network contributes about the same at both time points, though as noted previously, less than in the goal achievement model.

Chapter 7 - Conclusions

The Quick Response Study was an ambitious project which sought to examine objectively the claims put forward by proponents of Quick Response. The project was not successful because too few subjects were enrolled for reliable use of statistical techniques. Having said that, it was possible to learn some lessons by analyzing what went wrong, and even with small numbers, to observe some trends.

7.1 Inadequate enrolment of subjects: this is discussed in Chapter 4 (pp. 74 - 77.) Several factors were outside the control of the investigators: reduction in the number of hospital beds; staff anxiety over the proposed reorganization of the St. John's hospitals; and identification of potential subjects, and their referral to the study. In spite of the efforts of the investigators to inform staff and to enlist their support, most ER staff were too busy to take an active role in identifying patients who met QR criteria. It is possible that had QR nurses been recruited from the ranks of ER nurses at each hospital, they would have felt part of the QR team, and thus taken more responsibility for identifying and bringing appropriate patients to the attention of ER physicians. This illustrates the importance of bringing on side people who are vital links in the study process.

7.2 Study hypothesis: the primary hypothesis was that QR will cost less than SC while proving as effective, in terms of patient health, caregiver burden, and patient and caregiver satisfaction. A specific aim was to reduce inpatient days in the intervention (QR) group by comparison with the SC group by at least 25%.

7.3 Results: randomization worked well by age, and fairly well in terms of diagnosis. No differences, clinical or statistical, were evident between the two randomized groups in terms of Barthel score reflecting ADL abilities prior to or at the time of the ER visit, or in initial SF-36 or SIP scores. As anticipated, there was a wide range of scores for all questionnaires. No statistically significant difference between the two groups was observed in any questionnaire score, either at two weeks or at three months, but unfortunately, numbers were too low to conclude that no difference existed. Similarly, it was not possible to draw any conclusions concerning the efficacy of QR in relieving caregiver stress or affecting caregiver health. Although the two randomized subject groups appeared to be comparable, there was a baseline difference between their caregivers, as measured by perception of stress.

Considering hospital admissions to acute care beds occuring in the first two weeks (the period of QR intervention), the specific target of a 25% reduction of inpatient days relative to the size of each group was clearly met by the QR group. Over the three-month period, the group which had received QR support over the first two weeks still spent far fewer days in hospital per subject, although when non-randomized QR subjects were included, the difference was not significant.

Comparison of mean and median costs shows that Community Health costs for SC subjects were half that for randomized QR subjects. However, the overall cost of care for SC was almost twice (1.9 and 1.8) that of QR subjects. Impressive as this may seem, these differences were not statistically significant, at $p > 0.1$ with respect to Community Health costs, and $P > 0.5$ for total costs (Wilcoxon rank sum test).

7.4 ***Caveat***: the major problem with this study was the low number of subjects enrolled. Although randomization worked reasonably well, it is difficult to be certain that such small groups are comparable in characteristics, recognized and unrecognized, which may affect the outcome. Larger numbers would have afforded a greater degree of certainty that differences in outcome between the two groups resulted from the different intervention rather than baseline differences. Logistic regression identified three variables as predictors both of failure to attain goals, and also of poor two-week and three-month outcomes: greater age, lack of a supportive network of family and friends, and presence of one or more unfavourable diagnoses. The age difference between the two study groups was negligible. In terms of diagnosis, 7/20 SC subjects and 6/22 QR subjects had one or more adverse diagnoses. There was some difference in support network: 6/20 SC subjects and 2/22 subjects lacked support. Although this difference was not statistically significant, it may act as a confounding factor in the analysis, by increasing the likelihood of poorer outcome in the SC group. This weakens the findings of a difference between the SC and QR groups with respect to hospital days and costs. Study group (SC or QR, non-randomized and randomized) was not identified as an important variable in the regression analysis of goal achievement or health outcome.

An important question is whether the study sample is representative of the population from which it is drawn. This population is defined in the inclusion criteria, which were clearly articulated (see Appendix, page 127); about halfway through the Pilot Phase, the age criterion was modified to include all adults over the age of 18 years. From the date of this change, 28% of subjects were younger than 60.

The QR programs described in Chapter 2 were all targeted primarily at the frail elderly, but four of the six programs also accepted younger adults when appropriate. This pragmatic approach to identifying those who can benefit makes sense, and suggests that our definition of this population accords with that of many health care agencies. Potential subjects who were not enrolled were about 5 years older than enrolled subjects, (Student's t-test $p > 0.5$), but otherwise were similar.

7.5 Should a QR program be implemented? Although the striking reduction in hospital days might initially lead one to recommend the introduction of a QR team in the St. John's region, further consideration of the data does not entirely support this. It is possible that a much larger study might demonstrate an advantage, in terms of overall cost, to having a QR team, but given the rapid response time of routine Community Health professional nursing services (generally within 24 hours of the referral being received by telephone or fax, and often within hours if the case is urgent) it might not be possible to realize any savings, particularly if administrative costs of providing the extra service are included. A further consideration is that given the constant pressure to stay within budgetary restrictions, no community agency is willing to take on an additional commitment without clear assurance that it will be fully funded. Although a QR system might prove to be in the interest of the health care system as a whole, a directive with funding attached would have to be issued from the Department of Health for such a program to be implemented.

7.6 Opportunities for Improvement in Delivery of Care: one aspect of QR which appeared to be particularly important was the speedy provision of homemaking services. Several SC subjects who had to be admitted during the first two weeks would possibly have been able to remain at home given immediate non-professional home support, but either this was not recognized or there was no mechanism in the routine system for providing such service in a timely fashion. Admission to hospital is a costly response if provision of short-term home support could allow these individuals to remain at home. The Barthel questionnaire might be helpful in this context to give ER staff and family physicians a clearer idea of the functional abilities of patients and thus a guide to how much assistance they will need. This consideration is commonly neglected or underestimated by physicians.[47] Compounding the problem of sudden loss of self-sufficiency in ADL may be the inability or lack of family support to cope with the new situation. It would therefore be useful if a brief assessment of the strength of the available support network were combined with the Barthel questionnaire in the ER or physician's office: this would not take more than five minutes, and could be documented by the ER nurse before the discharge orders are written, or by the family practice nurse before the patient is seen by the family doctor. Where appropriate, follow-up could be provided by Community Health staff.

The ER management of people with a multiplicity of chronic problems, often with an acute problem superimposed, is challenging. They present ostensibly with an acute problem, perhaps a fall or urinary tract problem, but the underlying condition requires more careful evaluation than is possible in the ER. Patients with cognitive

impairment sometimes present in this way. To treat the presenting complaint without addressing underlying problems is an incomplete response, but the ER is not the ideal place to make long-term decisions concerning future management. More time and a quieter atmosphere than the ER can provide are necessary to assess the patient fully and to discuss options with the patient and family. The acute problem should be managed expeditiously, and the patient should be referred by the ER or consultant physician for in-depth assessment by the family physician or in some cases by a geriatrician and multidisciplinary team.

7.7 Research Opportunities: a useful research project would be to analyse and classify the problems of all adults who come to the ER having fallen: to record the diagnosis or diagnoses, prognosis and management; to record any documentation of ADL at the time of their visit, and any need for assistance or assistive equipment (and whether such was requested and put in place); and to record any subsequent visits to the ER relating to the same incident. What proportion of these patients had multiple active medical diagnoses? Did they receive any further assessment of their problems after the ER visit?

Another interesting project could be the analysis of all admissions to hospital through the ER, to assess the percentage (if any) who could have been cared for at home, given adequate support. Is there still a group of people for whom a version of QR could be helpful in avoiding admission? Findings could be compared with data from the 1992 QR feasibility study (see page 20). If such a group is identified, questions to be answered by chart review are essentially the same as noted above

for patients who had fallen, plus: what was available in hospital that could not have been provided at home?

Community care offers many opportunities for research. Involving staff by asking them to suggest questions would generate many good ideas; those who expressed interest would enjoy the learning process, and the exercise would contribute greatly to raising awareness of determinants of good care. With a clearly focused question, and careful recording of data, evaluation of care can be both interesting and productive. For this to work well, support from senior management, in the form of some protected time for research activities, would be essential.

A recent meta-analysis[48] of twenty studies (excluding studies dealing with children and psychiatric patients) examined the impact of Home Care (not Quick Response) on days in hospital. The conclusion was that Home Care significantly reduces the number of days spent in acute hospital beds, by 2.5 - 6 days per 180 days, effect size -.16 to -.38, where effect size = mean (intervention group) minus mean (controls) divided by pooled standard deviation. Although the authors examined 327 papers on this subject, only 20 met their criteria for inclusion. These criteria included using a control group or a pre-post design; comparison of home care with customary care; clear reporting of actual hospital days and/or cost, with means and standard deviations. The authors made a plea to home care agencies that when they implement new programs, they do so with a view to subsequent evaluation, by using precise definitions, by using randomized comparison groups and by clear reporting of patient health status and the relevant diagnoses, .

It is clear that health care in the community will play an increasing role, with hospital care reserved primarily for investigation and acute interventions. It is imperative that evaluation of care be an integral part of planning, not a cosmetic afterthought. This requires that before a new program is implemented, careful consideration be given to its objectives, how it will be evaluated and what data gathering will be necessary. Chapter 2 describes some QR programs which were put in place with enthusiastic but superficial evaluation. The St. John's QR project described here was not successful because of inadequate numbers, but it illustrates some of the important considerations in program evaluation. Research is an essential component of health care, and community health agencies are in an excellent position to incorporate a research approach into evaluation of all their services.

References

1. Clinch J. Letter to Edward Jenner, January 25, 1802, cited with other correspondence relevant to the introduction of vaccination against smallpox in Newfoundland. In: Roberts KB. Smallpox: an historic disease. Memorial University of Newfoundland, 1979.

2. Anspach, LA. A history of the island of Newfoundland. 2nd ed. London: Sherwood, Gilbert and Piper, 1827: 258.

3. Newfoundland Acts of Assembly: 3, William IV c.11, 1833.

4. Royal commission on health services report, Vol. 1, Ottawa: Queen's Printer, 1964.

5. Inouye SK, Wagner R. *et al.* A predictive index for functional decline in hospitalized elderly medical patients. J Gen Intern Med 1993; 8: 645 - 652.

6. Hirsch CH, Sommers L *et al.* The natural history of functional morbidity in hospitalized older patients. J Am Geriatr Soc 1990; 38: 1296 - 1303.

7. Inouye S. and Charpentier PA. Precipitating factors for delirium in hospitalized elderly persons. JAMA 1996; 275: 852 - 857.

8. Dawson JI and Critchley L. Community-hospital partnerships: the quick response team. J Nurs Adm 1992; 22(11): 33 - 39.

9. Boyack VJ and Bucknum AE. The quick response team: a pilot project. Soc Work Health Care 1991; 16(2): 55 - 68.

10. Darby PW. Quick response teams: a new approach in utilization management. Leadership Health Serv 1992; 1(5): 27 - 31.

11. Freeman M. Quick response programs: effective management of a population in crisis. Leadership Health Serv 1994; 3(50): 36 - 39.

12. Rajacich DL and Cameron S. Preventing admissions of seniors into the emergency department. J Gerontol Nurs 1995; 21(10): 36 - 40.

13. Williams G and Smith G. Assessing the role of home care in the deferral of hospital admissions through the emergency department. Wellington-Dufferin-Guelph Home Care Program, 1991.

14. Bolton C, Brazil K, *et al.* Kingston, Frontenac, Lennox and Addington: Quick Response Service, final report, February 1996.

15. Report of the Ottawa-Carleton Quick Response Service, November 1995.

16. Merbitz C, Morris J and Grip JC. Ordinal scales and foundations of misinference. Arch Phys Med Rehabil 1989; 70: 308 - 312.

17. Stucki G, Daltroy L, Katz J *et al.* Interpretation of change scores in ordinal clinical scales and health status measures. J Clin Epidemiol 1996; 49(7): 711 - 717.

18. Wright BD and Linacre JM. Observations are always ordinal; measurements, however, must be interval. Arch Phys Med Rehabil 1989; 70: 857 - 860.

19. Silverstein B, Fisher WP, Kilgore KM *et al.* Applying psychometric criteria to functional assessment in medical rehabilitation. II. Defining interval measures. Arch Phys Med Rehabil 1992; 73: 507 - 518.

20. Fortinsky RH, Granger CV and Seltzer GB. The use of functional assessment in understanding home care needs. Med Care 1981; 19(5): 489 - 497.

21. Mahoney FI and Barthel DW. Functional evaluation: the Barthel Index. Md State Med J 1965; 14: 61 - 65.

22. Granger CV, Albrecht GL and Hamilton BB. Outcome of comprehensive medical rehabilitation by PULSES profile and the Barthel Index. Arch Phys Med Rehabil 1979; 60: 145 - 154.

23. Bergner M, Bobbitt RA *et al.* The Sickness Impact Profile: development and final revision of a health status measure. Med Care 1981; 19: 787 - 805.

24. Deyo RA, Inui TS *et al.* Measuring functional outcomes in chronic disease: a comparison of traditional scales and a self-administered health status questionnaire in patients with rheumatoid arthritis. Med Care 1983; 21(2): 180 - 192.

25. Beaton DE, Hogg-Johnson S and Bombardier C. Evaluating changes in health status: reliability and responsiveness of five generic health status measures in workers with musculoskeletal disorders. J Clin Epidemiol 1997; 50(1): 79 - 93.

26. MacKenzie CR, Charlson ME *et al.* Can the Sickness Impact Profile measure change? An example of scale assessment. J Chron Dis 1986; 39(6): 429 - 438.

27. Katz JN, Larson MG *et al.*. Comparative measurement sensitivity of short and longer health status instruments. Med Care 1992; 30: 917 - 925.

28. Ware JE and Sherbourne CD. The SF-36 health status survey. 1. Conceptual framework and item selection. Med Care 1992; 30(6): 473 - 483.

29. McHorney CA, Ware JE *et al.* The MOS 36-item short-form health survey (SF-36): III: tests of data quality, scaling assumptions and reliability across diverse patient groups. Med Care 1994; 32(1): 40 - 66.

30. Brazier JE, Harper R *et al.* Validating the SF-36 health survey questionnaire: new outcome measure for primary care. BMJ 1992; 305: 160 - 164.

31. McHorney CA, Ware JE and Raczek AE. The MOS 36-item short form health status survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993; 31(3): 247 - 263.

32. Stewart AL, Greenfield S *et al.* Functional status and well-being of patients with chronic conditions. JAMA 1989; 262: 907 - 913.

33. Andresen ME, Patrick DL *et al.* Comparing the performance of health status measures for healthy older adults. J Am Geriatr Soc 1995; 43: 1030 - 1034.

34. Fryback DG, Dasbach EJ *et al.* The Beaver Dam health outcomes study: initial catalog of health-state quality factors. Med Decis Making 1993; 13: 89 - 102.

35. McHorney CA. Measuring and monitoring general health status in elderly persons: practical and methodological issues in using the SF-36 health survey. Gerontologist 1996; 36(5): 571 - 583.

36. Greene JG, Smith R *et al.* Measuring behavioural disturbance of elderly demented patients in the community and its effects on relatives: a factor analytic study. Age Ageing 1982; 11: 121 - 126.

37. Draper BM, Poulos CJ *et al.* A comparison of caregivers for stroke and dementia victims. J Am Geriatr Soc 1992; 40: 896 - 901.

38. Pattie AH and Gilleard C. Manuals of the Clifton assessment procedures for the elderly. Sevenoaks, Kent: Hodder and Stoughton, 1979, cited in reference 36.

39. Lawton MP and Brodie EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969; 9: 179 - 186.

40. Eagles JM, Craig A *et al.* The psychological well-being of supporters of the demented elderly. Br J Psychiatry 1987; 150: 293 - 298.

41. Goldberg DP. The detection of psychiatric illness by questionnaire. Maudsley Monograph No. 21. London: Oxford University Press, 1972.

42. Goldberg DP. Manual of the General Health Questionnaire. Windsor, England: National Foundation for Educational Research, 1978.

43. Goodchild ME and Duncan-Jones P. Chronicity and the General Health Questionnaire. B J Psychiatry 1985; 146: 55 - 61.

44. Andrews FM and Withey SB. Social indicators of well-being. Americans' perception of life quality. New York: Plenum Press, 1976.

45. Freiman JA, Chalmers TC *et al.* The importance of beta, the Type II error and sample size in the design and interpretation of the randomized control trial. N Engl J Med 1978; 299: 690 - 694.

46. Julious SA, George S and Campbell MJ. Sample sizes for studies using the short form 36 (SF-36). J Epidemiol Community Health 1995; 49: 642 - 644.

47. Calkins DR, Rubenstein LV *et al.* Failure of physicians to recognize functional disability in ambulatory patients. Ann Int Med 1991; 114: 451.

48. Hughes SL, Ulasevich A *et al.* Impact of home care on hospital days: a meta-analysis. Health Serv Res 1997; 32(4): 415 - 432.

APPENDIX

A. Eligibility Criteria

B. Box and Whisker Plots

C. Logistic Regression

A. Eligibility Criteria

1. Patients shall be eligible for inclusion in the study if they:

   1.1 present to the emergency room of the General Hospital, St. Clare's Mercy Hospital or the Salvation Army Grace General Hospital.

   1.2 live within the boundaries of the region covered by Community Health - St. John's Region, with the following exceptions: Bell Island and Bay Bulls Big Pond to St. Shotts on the Southern Shore. The Paradise area is limited to St. Thomas' Line to Topsail Road including Paradise up to the junction of Topsail Pond Road and Three Island Pond Road.

   1.3 are aged 60 years or older.

   1.4 are not affected by a medical or social condition that either (1) mandates immediate inpatient hospital care, or (2) is entirely treatable by services available within the emergency room, with or without further consultation.

   1.5 might benefit from the availability of "enhanced" care in the home, within four hours of discharge home, from Registered Nurses, Registered Nursing Assistants, Home Support Workers, Social Workers, Occupational Therapists and/or Physiotherapists.

   1.6 are eligible for coverage under the Newfoundland Provincial Health care plan (MCP).

2. Patients shall be excluded form the study if they do not meet the eligibility criteria or

   2.1 during the Evaluation Phase the patient or his or her caregiver refuses consent to take part in the randomized trial.

   2.2 they are affected by a medical condition that could render them dangerous to themselves or others as determined by the responsible physician in the emergency room or staff of the Quick Response Team.

   2.3 care in the home is not acceptable to the patient or his or her immediate caregivers.

   2.4 care in the home is not thought possible by the Assessment Nurse from the Quick Response Team because of the home circumstances.

   2.5 the nature or level of care required exceeds that available through the Quick Response Program, as determined by the Assessment Nurse.

B. Box and Whisker Plots

Box and whisker plots shown in Figures 4.5 - 4.9, 4.13 - 4.14, 6.4 - 6.5 and 6.10 - 6.11 were generated using SPSS 6.1 software (SPSS, Chicago, Illinois).

Boxes show the **median** value, and the **25th** and **75th percentiles**.

Whiskers extend to the lowest and highest scores, excluding outlying and extreme values.

Outlying values (o) are defined as falling more than 1.5 box lengths beyond the 25th or 75th percentile box end.

Extreme values (*) lie more than 3 box lengths beyond either box end.

C. **Logistic Regression** (taken from SPSS Advanced Statistics 6.1, Norušis M.J. SPSS, Chicago, Illinois, 1994)

Coefficients are estimated using the maximum likelihood method, *i.e.* the coefficients selected make the observed results most "likely". The logistic regression model is non-linear: when the probability of an event (always between 0 and 1) is plotted against values of Z, the curve is S-shaped, closely resembling the curve of the cumulative probability of the normal distribution.

$$\text{Probability (event)} = \frac{e^{(B_0 + B_1 X_1)}}{1 + e^{(B_0 + B_1 X_1)}}$$

$$\text{or} \qquad \frac{1}{1 + e^{-(B_0 + B_1 X_1)}}$$

where B_0 and B_1 are coefficients estimated from the data,

X is the independent variable and e is the base of the natural logarithm, ~ 2.718.

For more than one independent variable, the model can be written as

$$\text{Probability (event)} = \frac{1}{1 + e^{-z}}$$

where Z is the linear combination $B_0 + B_1X_1 + B_2X_2 \ldots B_pX_p$

The probability of an event not occurring, Prob(no event) is estimated as 1 - Prob(event).

Categorical variables: values of independent variables are recoded by creating a new set of variables which correspond in some way to the original categories. In the regression analyses performed here, support and adverse diagnosis were coded 0.5 when present, - 0.5 when absent. When these new values are inserted into the equation and multiplied by the appropriate coefficient, they contribute the calculated weighting to the regression model which predicts an outcome or class depending on the presence or absence of these independent variables.

Forward stepwise selection is the method used to enter variables in the model. The model starts with a constant, and at each step, a variable is entered, (starting with the variable with the smallest significance level for its score statistic, provided it is less than the cutoff value of 0.05). All variables entered into the model are then examined in turn and if the likelihood ratio (the change in log likelihood when a given variable is deleted) for any variable exceeds the cutoff value (set at 0.1) it is removed.

There are several ways of describing how well the model fits the data: one of these involves **likelihood**, which is defined as the probability of the observed results, given the parameter estimates. It is a small number, less than unity, and minus twice the log of the likelihood (**-2LL**) is used as the measure. A good model results in high likelihood of the observed results, which translates to a small value for -2LL. If a model fits perfectly, the likelihood is 1; -2 times the log likelihood is 0. Another measure is the **goodness-of-fit statistic** which compares the observed probabilities to those predicted by the model.

The **classification table** compares prediction of cases by the model with observed outcomes, showing how many (by percentage) of those with the outcome of interest were correctly predicted, and similarly, how many of those who did not develop the outcome of interest had been correctly predicted as unlikely to develop it.

Partial correlation. The contribution of each variable depends on the other variables in the model; the statistic which expresses the partial correlation between the dependent variable and an independent variable (given the presence of the other variables in the model) is the **R** statistic, whose value can range from -1 to +1. A positive value for R indicates that as the variable increases in value, so does the likelihood of the event occurring. Conversely, if R is negative, the opposite is true, *i.e.* with an increase in value of the variable, the likelihood of the event decreases. The lower the value of R, the smaller the partial contribution to the model made by that variable.

Interpreting the regression coefficients. The logistic model may be written in terms of the log of the odds of an event occurring, known as the **logit**:

$$\log \frac{\text{Prob(event)}}{\text{Prob(no event)}} = B_0 + B_1X_1 + \ldots . B_pX_p$$

The regression coefficient can be interpreted as the change in the log odds associated with a one-unit change in the independent variable. For example, in the example of prediction of failure to achieve the goal at two weeks (page 108), the coefficient for age is .0857. Provided that the values of other independent variables remain the same, this means that for an increase in age of 1 year, the log odds of failure increase by .0857.

For data used to predict achievement at two weeks, the "event" selected by the computer was failure to achieve or partly achieve the goal. Probability of failure = $\frac{1}{1 + e^{-z}}$

where $Z = B_0 + B_1(\text{support}) + B_2(\text{diagnosis}) + B_3(\text{age})$

where B_0 = a constant, - 6.1259, B_1 = - 4.1546, B_2 = 2.9489, and B_3 = .0857

For the categorical variables:

Support positive (supportive network of family and/or friends)	enter	0.5,
Support negative (lack of supporting network)	enter	- 0.5
Adverse medical diagnosis (presenting or concurrent)	enter	0.5
No adverse medical diagnosis (presenting or concurrent)	enter	- 0.5

If the probability estimate is < 0.5, the case is predicted to be "no event" and if the estimate > 0.5, the prediction will be for an "event".

Thus for an individual aged 91, without support and with an adverse diagnosis,

$$Z = -6.1259 - 4.1546 \times -0.5 + 2.9489 \times 0.5 + .0857 \times 91$$

$$= -6.1259 + 2.0773 + 1.4745 + 7.7987$$

$$= 5.2246$$

$$\text{Then prob(failure)} = \frac{1}{1 + e^{-5.2246}}$$

$$= \frac{1}{1 + 0.0054} = \frac{1}{1.0054} = 0.9946$$

Predicted probability is .9946, (> 0.5) therefore the event "failure" is predicted to occur, and this case is assigned to the "will not achieve goal at two weeks" group.

Returning to the interpretation of the regression coefficients (see pp. 131,132) the log of the odds of an event occurring, known as the **logit**, is expressed as

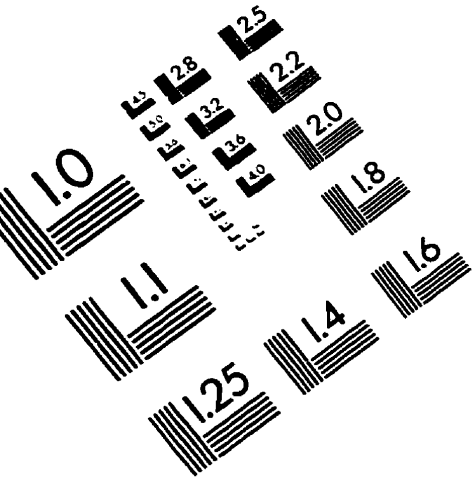
$$\log \frac{\text{Prob(event)}}{\text{Prob(no event)}} = B_0 + B_1X_1 + \ldots B_pX_p$$

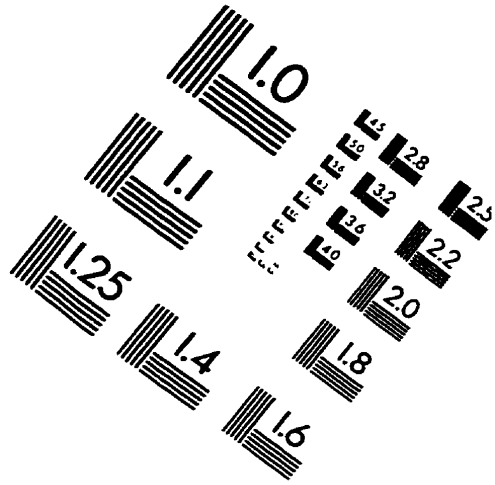
By taking the antilog of both sides, the equation may also be written in terms of the odds:

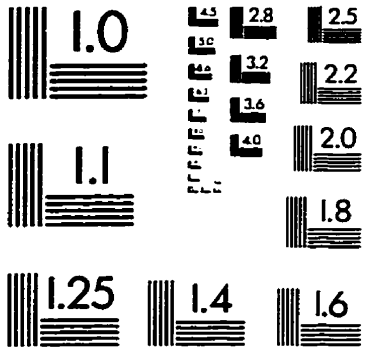
$$\frac{\text{Prob(event)}}{\text{Prob(no event)}} = e^{B_0 + B_1X_1 + \ldots B_pX_p}$$

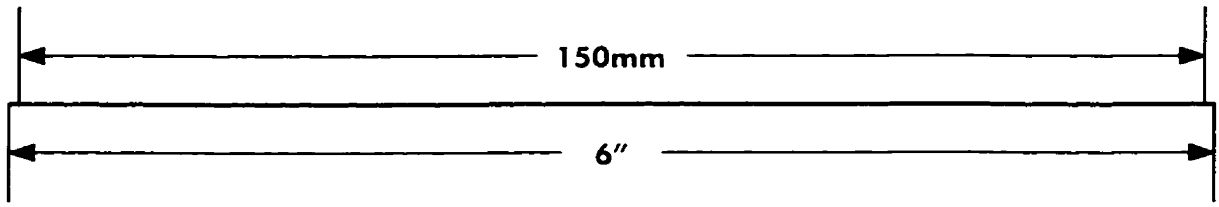
The **exponent of B** then shows that when age increases by 1 year, the **odds** of failure to achieve are increased by a **factor** of 1.0894. Similarly, if the "adverse diagnosis" status changes, the odds of failure change by a factor of 19.0846.

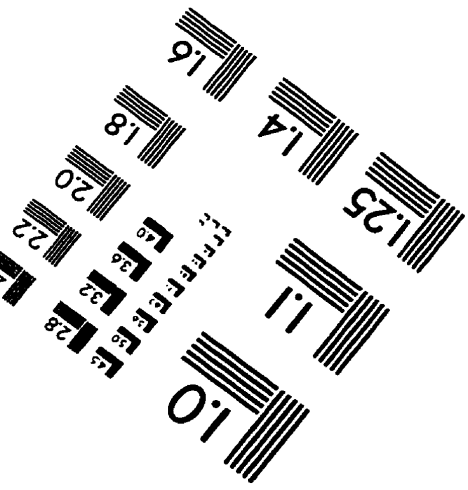
IMAGE EVALUATION
TEST TARGET (QA-3)

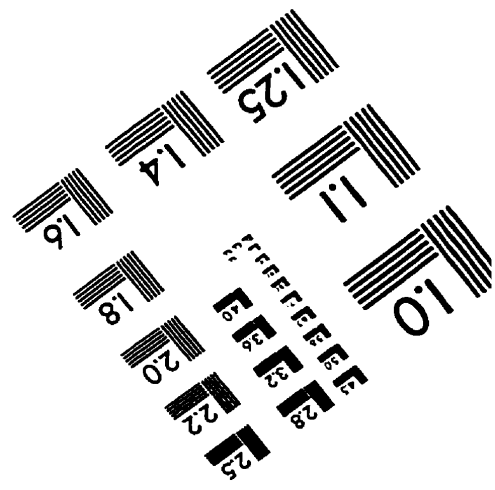












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989