

# Building an evidence base for community health: a review of the quality of program evaluations

Gwyn M Jolley, Angela P Lawless, Fran E Baum, Catherine J Hurley and Denise Fry

## Abstract

An assessment of the quality of program evaluations conducted in South Australian community health services investigated how effective evaluation reporting is in producing an evidence base for community health. Evaluation reports were assessed by a team of reviewers. Practitioner workshops allowed an understanding of the uses of evaluation and what promotes or acts as a barrier to undertaking evaluations.

Community health services do undertake a good deal of evaluation. However, reports were not generally explicit in dealing with the principles that underpin community health. Few engaged with program theory or rationale. Typically, reports were of short-term projects with uncertain futures so there may seem little point in considering issues of long-term health outcomes and transferability to other settings. The most important issue from our study is the lack of investment in applied health services research of the sort that will be required to produce the evidence for practice that policy makers desire. The current lack of evidence for community health reflects failure of the system to invest in research and evaluation that is adequately resourced and designed for complex community settings.

*Aust Health Rev* 2007; 31(4): 603–610

**Gwyn M Jolley**, MSC(PHC), Senior Research Officer  
**Angela P Lawless**, BAppSci(SpeechPathology), GradDipPHC,  
 Senior Research Officer

**Fran E Baum**, PhD, Director and Professor of Public Health

**Catherine J Hurley**, BA, BSocAdmin,  
 GradCertPHC(Research and Evaluation), Senior Research  
 Officer

South Australian Community Health Research Unit,  
 Department of Public Health, Flinders University, Adelaide, SA.

**Denise Fry**, BA(Hons), Walk To School Project Officer  
 Health Promotion Service, Sydney South West Area Health  
 Service, Sydney, NSW.

Correspondence: Ms Gwyn M Jolley, South Australian  
 Community Health Research Unit, Department of Public Health,  
 Flinders University, G Block FMC Flats, Flinders Drive, Bedford  
 Park, Adelaide, SA 5042. [gwyn.jolley@flinders.edu.au](mailto:gwyn.jolley@flinders.edu.au)

## What is known about the topic?

Program evaluations are often conducted for community health initiatives, but there has been little information on the quality of the completed evaluations.

## What does this paper add?

This paper presents the results of a review of program evaluations conducted in South Australian community health services. Although there were a large number of evaluations completed, most were internal and did not provide useful information for policy and planning decisions.

## What are the implications for practitioners?

The authors suggest a need for investment in health services research to improve the quality of program evaluations for decision making.

**THIS PAPER DESCRIBES** a review of the quality of program evaluations conducted in South Australian community health services and considers how effective these evaluation reports are in terms of producing an evidence base for the work of these services. It also reports on the attitudes of services toward evaluation. Community health in this paper describes the state government-funded primary health care sector that provides comprehensive primary health care services not targeted at a particular population group. Services are mainly non-medical and multidisciplinary, with a range of strategies and an emphasis on health promotion and illness prevention.

Over the last decade there has been an increasing interest in evidence-based medicine (EBM) and the application of evidence-based principles to other areas of health practice and policy making. For example, a MEDLINE search for evidence-based medicine revealed one citation in 1992 but more than 13 000 in 2004.<sup>1</sup> EBM mainly uses systematic reviews of randomised and other controlled trials to assess and synthesise evidence about the effectiveness of interven-

tions. Increasingly, primary health care and health promotion programs are being called upon to produce similar evidence of effectiveness. While much has been written about the (un)suitability of direct translation of EBM methods, given the complexity of primary health care interventions and settings<sup>2,3,4</sup> the pressure to develop an evidence base for primary health care practice and policy continues to grow.

There is value in primary health care policy makers and providers developing evidence on the effectiveness of their practice. Concern has been expressed that the progress of community health is impeded by the lack of documented evidence for practice and programs and a belief that it is not possible to produce robust evidence such as that from randomised controlled trials. This means that community health is often disadvantaged when arguing for funding, particularly when competing with hospitals and acute care services where there is less demand for evidence as a basis for funding.<sup>5</sup>

This is not to say that community health services do not have a strong commitment to research and evaluation. In fact, SA community health service programs routinely include some form of evaluation.<sup>6</sup> Given the small proportion of the health dollar received by community health in Australia nationally (\$3.1 billion or 4.8% of total recurrent expenditure in 2002–03),<sup>7</sup> it compares favourably with other parts of the health sector in terms of evaluation practice.<sup>5</sup> To meet the calls for evidence, however, an approach to evaluation that moves beyond evaluation for internal organisational purposes to one which provides useful evidence for the development and improvement of community health practice is needed.

Community health practitioners and policy makers involved in the development and implementation of services and programs in areas as diverse as mental health, child development, violence intervention, physical activity and healthy ageing need a robust evidence base and resources with which to produce this. While the dangers of deciding on the wrong treatment in a clinical context seem obvious, the dangers of implementing the wrong program or policy in

response to a community health issue may be equally far reaching. Some apparently well intentioned interventions have had adverse unintended consequences, such as a bicycle education program which actually increased the risk of injury,<sup>8</sup> and an eating disorders prevention program that had short-term success but at 6-months showed a return to baseline levels for eating disorders and an unwelcome increase in dietary restraint.<sup>9</sup> There is also evidence to suggest that health promotion messages are taken differentially by different population groups. For example, since the promotion of folate and voluntary fortification of food, there has been a 30% fall in neural tube defects in Western Australia. However, there has been no reduction in rates in the Aboriginal population, and neural tube defects in Aboriginal infants are almost twice as common compared with non-Aboriginal infants.<sup>10</sup>

Thus it is possible for such programs to have an unintended effect of increasing inequities. It is important that providers can be confident that their programs are beneficial to participants and the wider community, that practitioners have a good understanding of what interventions are effective, and why and what may cause harm. This type of evidence is important to convince decision makers to fund programs, to convince policy makers to extend successful programs and to inform decision making about the opportunity costs involved in choosing one program over another.

### **Towards an evidence base**

The first phase of this work is reported in *Investing in community health — finding the evidence for effectiveness*,<sup>5</sup> which identified three major challenges in establishing an evidence base for community health: the difficulties inherent in attributing program outcomes to a range of interventions; the complexity of the community-based setting; and the danger of ease of measurement driving the intervention. Four means by which community health programs can be judged were examined: economic evaluation, use of routine databases, systematic reviews and performance

**I Mean score for each review question**

<b>Question</b>	<b>Mean score (1–5)</b>
Q1.1 Does the evaluation provide a clear description of the program goals/aims/expected outcomes?	3.462
Q1.2 Does the evaluation provide a clear description of the intervention/program and the processes used in it?	3.532
Q1.3 What evidence is presented that shows why the intervention is expected to lead to better health outcomes? (ie, Is the program logic well articulated?)	2.957
Q1.4 Does the evaluation consider issues of equity and produce evidence for the ways in which the intervention is working towards both a) equity of access to services and b) equity in health outcomes?	2.968
Q1.5 Does the evaluation document ways in which the broader implications of a health issue are considered through the intervention? Are attempts made to tackle “up stream” causes of the problem?	2.828
Q1.6 Does the evaluation discuss to what extent and how effectively the intervention involves community participants?	2.978
Q1.7 Does the evaluation discuss to what extent and how effectively the intervention involves other groups and agencies?	2.763
Q1.8 Does the evaluation document unintended aspects of the intervention?	2.548
Q1.9 Does the evaluation report on achievement of program objectives/expected outcomes? Are immediate and intermediate outcomes reported?	3.274
Q1.10 Does the evaluation discuss the likelihood of achieving longer term health outcomes?	2.349
Q1.11 Does the evaluation report on transferability of the intervention?	2.339
Q1.12 Does the evaluation report on sustainability of the outcomes?	2.309
Q2.1 Does the evaluation provide a sound justification for the evaluation method and acknowledgement of limitations of the method chosen?	2.677
Q2.2 Does the evaluation use a representative sampling method for those consulted as part of the evaluation?	2.629
Q2.3 Does the evaluation provide an adequate description of the context of intervention?	3.016
Q2.4 Does the evaluation provide evidence of data quality?	2.715

indicators. Each of these techniques was reviewed to determine the contribution they make to assessing the effectiveness of community health services. The resulting report concluded with a discussion of the way forward to a more systematic approach to assessing the performance of community health services. This paper presents a review of evaluation reporting, not of the interventions themselves.

## Methods

From previous experience with the community health sector, the research team were aware that evaluations are inevitably small scale and rely

heavily on qualitative methods. Qualitative systematic review methodology is underdeveloped in comparison to statistical meta-analysis and systematic review, and there is no agreed method for assessing the quality of qualitative studies.<sup>11</sup> A review framework was proposed by the research team to assess the quality of reporting on planning, program logic and evaluation in community health services. The framework also needed to take account of the importance of the core values underpinning comprehensive primary health care, especially participation, equity and recognition of the social determinants of health.<sup>5</sup> These core values both define and strengthen primary health care delivery in SA community health

## 2 Sample questions from the review protocol

Q. 1.3 Does the evaluation provide a program logic?

*Explaining the logic behind an intervention is important*

What evidence is presented that shows why the intervention is expected to lead to better health outcomes? (ie, Is the program logic well articulated?)

1	2	3	4	5	N/A
not met	minimally met	partially met	largely met	fully met	justify
Comments					

Q 2.1. Does the evaluation provide a sound justification for the evaluation method and acknowledgement of limitations of the method chosen?

*The evaluation methods used should be described fully with sufficient details for a reviewer to make a judgement about their applicability to the particular evaluation.*

1	2	3	4	5	N/A
not met	minimally met	partially met	largely met	fully met	justify
Comments					

services and contribute to the difference in approach when compared with the medical model of health care.

The review protocol, based on previous work on synthesis of qualitative data<sup>12</sup> and public health interventions<sup>13</sup> was developed in consultation with a reference group comprising the research team, community health practitioners and state government policy officers. There were twelve questions about the description of the intervention and four questions about the evaluation methodology (see Box 1 for a list of question topics). Many of the questions had supplementary questions to guide the reviewer (see Box 2). The full review protocol can be viewed at <<http://som.flinders.edu.au/FUSA/SACHRU/Research/reviewtrialv2.doc>>

All evaluation reports from the five community/women's health services in the metropolitan region (1999–2002) were identified and collected. Inclusion/exclusion criteria resulted in a set of 93 reports for review. Reports were included only if the evaluation was formally documented and a metropolitan community health service was a key player in the activity/program. Reports also needed to contain, at a

minimum, a description of the intervention, a description of the evaluation method and a report of the findings.

The review team consisted of three researchers, a practitioner from each community health service and an interstate consultant with considerable experience in primary health care research and evaluation. Training sessions were held to maximise consistency and to finalise the review protocol questions.

Each report was independently reviewed twice: once by the interstate reviewer and once by a member of the SA review team. The extent to which each report met the descriptor was scored from 1 (not met) to 5 (fully met) (Box 2) and comments were invited for each question. Report and intervention characteristics, scores and comments were entered into SPSS version 11 (SPSS Inc, Chicago, Ill, USA) for collation and analysis.

At the request of the community health services, short workshops on evaluation at each of the participating community health services were conducted. The purpose of the workshops was to gain an understanding of the current uses of evaluation within services and the factors that promote or act as barriers to practitioners under-

### 3 Promoters and barriers to evaluation (in order of response frequency)

Promoters	Barriers
Skills and training	Not enough time/resources for evaluation
Culture of evaluation	Lack of evaluation culture
Evaluation process or structure	Not enough expertise within organisation to do evaluation
Evaluation used to make a difference	Evaluation results aren't used
Access to expertise and support	External evaluation too expensive
Appropriate data systems and evaluation tools	Evaluation not seen as relevant/appropriate to work
Consistent framework	Evaluation is perceived as a threat to individual or program
Feedback and follow up	Don't know how to interpret evaluation findings

taking evaluation of their work. Six workshops took place with 127 participants in total. The process included a round table discussion about current uses of evaluation, a listing of individual, organisational and system level evaluation promoters and scoring a list of potential barriers. After identification of the top three barriers for each group, there was discussion about how these barriers might be addressed.

## Results and discussion

This was the first time that SA community health program evaluations had been subject to assessment of quality against a common set of criteria. The study revealed a large amount of varied and innovative program activity within community health services, and similarly diverse evaluation practice and reporting styles. It is important to understand that most of these evaluation reports had been written for an internal audience, and thus information that would have been made explicit in documents intended for a wider audience were consequently sometimes omitted from the evaluation report.

The total possible score for each evaluation report was 160. Assuming equal weighting across the 16 questions, scores ranged from 57 (36%) to 145 (91%) with a mean of 89 (56%). This represents the high end of “minimally met”. Summary scores for each question are listed in Box 1. Description of the goals and the intervention scored most highly; questions relating to

long-term outcomes, transferability and sustainability scored lowest.

The wide range of total scores and the consistency between reviewers suggests the review tool was robust and scores were not just a reflection of individual interpretation. The two reviewers' scores differed by 2 or more for 4.3%–4.4% of scores across the questions. Sustainability and sampling questions showed most frequent difference between reviewers.

## Description of programs

Most reports contained a clear description of the program goal and strategies. Problem definition and information about how the problem came to be identified were less clearly articulated, and low scoring reports typically lacked information about the intervention.

Describing the program logic and linking this to longer term health outcomes was generally poorly done. Given the intended audience for most reports was the health service itself, familiarity with the program and its development was probably assumed by the writers. Few reports tackled or discussed macro-level social or economic determinants of health or underlying causal issues.

Questions regarding three key features of community health practice — equity, community participation and collaboration — were included in the review tool. Reviewers found that these rarely figured in reports, despite the fact that primary health care principles are stated

as underpinning most programs and services. Again, the intended audience for reports may have influenced the way in which these issues were dealt with. For example, equity was not explicitly discussed even when the program was apparently designed with equity issues in mind, or equity issues were implicit in the focus of interventions, for example disadvantaged groups or geographical areas. Reviewers found various understandings of equity across disciplines and programs.

Community participation, while enshrined within service policy and strategic plans, also did not feature highly in the evaluation reports. Program participants were most often described as being involved by virtue of contributing to the evaluation or, to a lesser extent, to the planning stages. Participation was generally measured by attendance. Only a few reports reflected on the quality and effectiveness of community participation. A positive comment was:

Community-owned project and evaluation, involvement in structure, running and evaluation.

while a more typical comment was:

Not evident in report — didn't appear to be any community involvement in project beyond the participation in the forum.

Reporting on collaborative partnerships with other groups or agencies was similarly sparsely covered. A few reported comprehensively on the role of partnerships but others simply listed partners and did not describe the process of participation or reflect on the effectiveness of collaboration or report partners' views.

Reviewers were generally positive about reporting of achievement of objectives or immediate outcomes (for example, participants learned new skills around food and nutrition), although there were some concerns about the validity of data and findings. Longer term outcomes, such as actual changes to more healthy eating, were less often considered and seldom linked to established research. Evaluation reports generally did not address issues of the

potential for transferability and sustainability of the interventions.

### **Description of evaluation**

Nearly one-third of reports described only one method of data collection, usually participant feedback sheets. Another third reported using three or more methods. Many of the reports contained little or no justification of methods or limitations identified:

Sole evaluation method was feedback from parents in a questionnaire at the end of the group. No justification for this method or acknowledgement of limitations.

The question of the representativeness of the sample of those responding to the evaluation was generally not well reported. In most cases, where representativeness was covered, it was because the evaluation had included all people involved in the program in the evaluation. This was usually possible where the program was small. In many cases the response rates were not made clear or were left out altogether.

A few evaluations gave details of data analysis or identified more than one source of data. However, many evaluations failed to give sufficient detail in this area, particularly when it came to the analysis and presentation of data:

Results presented are unclear — not much about collection or analysis of data. Only one source of data is given.

### **Practitioner perspectives on evaluation**

Workshops with community health practitioners identified how evaluation was used within the organisation and the promoters and barriers to undertaking evaluation. Participants reported that evaluation was used for planning and improvement, accountability, validation and promotion of services and programs. Participants believed evaluation was more likely to be undertaken if it had a clear purpose and the findings were seen to be useful. Participants maintained that to establish an evaluation culture, an organisation should articulate the purposes for evaluation, establish a consistent framework, provide resource support and

structures, and access to expert help and guidance. Evaluation tools are needed that are relevant to the participants and the community, are flexible, qualitative and allow for creative methods of data collection and presentation.

The main promoters and barriers to evaluation identified by practitioners within the service are shown in Box 3. All workshops suggested “not enough time/resources” as the main barrier to evaluation. This was followed by “lack of evaluation culture” and “not enough expertise within organisation to do evaluation”. Much of the discussion regarding the time/resource barrier centred on the pressure to provide services and administrative and management responsibilities. This meant there was little time for reflection and evaluation.

## Conclusion

Our study suggests that community health services in SA do undertake a good deal of evaluation. The evaluation reports reviewed illustrate the enormous amount of innovative work being undertaken, in relation to some of the most complex issues and marginalised people in our society. Evidence from evaluations is used to inform planning processes and decision making and to describe programs and services to funders, bureaucracies and communities. Very few evaluations engaged with more fundamental theory or the underlying rationale for the program, even though some attention to this is usually required at the planning or funding submission stage. Most were internal evaluations, and a very small proportion were undertaken by external evaluators (usually for larger, grant-funded programs). Since the intended audience is mostly internal, it is reasonable to assume that knowledge of the program's rationale is assumed by the report writers. Further, when programs are limited by fixed funding and timeframes, there may seem to be little point in considering broader issues of long-term health outcomes, transferability to other settings, and so on. Typically, these evaluations were of short-term projects with uncertain futures. The writers of these reports were, for the most part, busy practitioners undertaking evaluation and report writing

with very little support in terms of time, resources or professional development.

A number of issues emerge from this study that must be addressed if community health services are to build evidence bases for their practice and programs that are convincing to funders. Firstly, systematic investment must be provided to support quality evaluations and their dissemination. With additional resources and greater expertise, evaluation and research will be able to develop and move from evaluations designed for mainly internal consumption to longer term research and evaluation with a focus on outcomes and program extension. The current investment in evaluation of community-based primary health services is very low when compared with other sectors, for example the General Practice Evaluation Program in the 1990s.

Secondly, practitioners in the workshops felt that their organisations did not have a culture supportive of evaluation despite the number and range of evaluation reports identified through the review process. The large number of evaluation reports was not reflected in subsequent use of findings. This suggests the need for more organisational commitment and support for workers undertaking evaluations. Organisations need to develop “learning cultures” that are demonstrated through organisational structures, processes and policies, for example, appropriate funding and in-kind support provided for evaluation and research activity. This culture would foster practitioners' ability to access and assess evidence and develop their research and evaluation skills and knowledge. Likewise publication and dissemination of evaluations must be facilitated and utilisation emphasised in order to contribute towards the broader evidence base.

Our review indicated that the evaluations were not generally explicit in dealing with the principles and strategies that underpin community health work. In particular, evaluations of equity, community participation and intersectoral collaboration, which are central to a community health approach, were mostly not well documented in the reports. For example, equity requirements were considered to be met if the program was targeted at a disadvantaged group; community

participation was frequently limited to consultation or opportunities to provide feedback; and partnerships were described but not assessed in terms of process or outcomes. Mechanisms need to be developed so that community health services can routinely articulate and evaluate these aspects of their programs.

### **Implications for research and practice**

A number of initiatives designed to enhance the capacity of community health services to undertake quality evaluations are being implemented as a result of this research. An evaluation and reporting template<sup>14</sup> was developed as part of the project in order to encourage consistency across reports and promote greater rigour. To encourage evaluation, tools for assessing partnerships, community participation and equity are being developed.

The most important issue to emerge from our study is the lack of investment in applied health services research of the sort that will be required to produce the evidence base for practice that services and policy makers desire. Most community health evaluation activity is not funded or is inadequately funded. The vast majority of research funding is directed at medical research rather than health systems research, a problem noted internationally, not just within Australia.<sup>15</sup> At times the push for “definitive evidence that programs work” appears daunting for community health practitioners, as producing evidence for their work is methodologically challenging and there are few resources to design and implement appropriate evaluations. Without adequate resourcing of, and commitment to, the development of high quality evaluation, reporting and dissemination, it will not be possible to produce an evidence base for community health programs that is comparable with that being established in the EBM world. The current lack of evidence for community health reflects failure of the system to invest in research and evaluation that is adequately resourced and designed for programs in complex community settings.

### **Acknowledgements**

South Australian metropolitan community health services provided access to evaluation reports. Community health

staff gave their time to sit on the reference group for the study and to participate as members of the review team.

### **Competing interests**

The authors declare that they have no competing interests.

### **References**

- 1 Strauss S. What's the E for in EBM? *BMJ* 2004; 328: 535-6.
- 2 Baum F. The new public health. 2nd edition. Melbourne: Oxford University Press, 2002.
- 3 Judge K. Testing evaluation to the limits: the cause of English health action zones. *J Health Serv Res Policy* 2000; 5: 3-5.
- 4 Salisbury C. Healthy living centres. *BMJ* 1999; 319: 1384-5.
- 5 Baum F. Introduction. In: Baum F, Duffy J, Jolley G, eds. Investing in community health — finding the evidence for effectiveness. Adelaide: South Australian Community Health Research Unit, 2003.
- 6 Jolley G, Baum F, Hurley C, Fry F. What works? A systematic review of the evidence for the effectiveness of community health in Adelaide. Adelaide: South Australian Community Health Research Unit, 2004.
- 7 Australian Institute for Health and Welfare. Health Expenditure Australia 2002-03. Canberra: AIHW, 2004. (AIHW cat. no. HWE27.)
- 8 Carlin J, Taylor P, Nolan T. School based bicycle safety education and bicycle injuries in children: a case-control study. *Inj Prev* 1998; 4: 22-7.
- 9 Carter J, Stewart A, Dunn V, Fairburn C. Primary prevention of eating disorders: might it do more harm than good? *Int J Eating Disorder* 1997; 22: 167-72.
- 10 Bower C, Eades S, Payne J, et al. Trends in neural tube defects in Western Australian Indigenous and non-Indigenous populations. *Paediatr Perinat Epidemiol* 2004; 18: 277-80.
- 11 Mays N, Pope C, Popay J. Systematically reviewing qualitative and quantitative evidence to inform management and policy-making in the health field. *J Health Serv Res Policy* 2005; 10 Suppl 1: 6-20.
- 12 Popay J, Rogers A, Williams G. Rationale and standards for systematic review of qualitative literature in health services research. *Qual Health Res* 1998; 8: 341-51.
- 13 Rychetnik L, Frommer M. A proposed schema for evaluating evidence on public health interventions version 3.1. Melbourne: National Public Health Partnership, 2001.
- 14 South Australian Community Health Research Unit. Are you writing an evaluation report? Evaluation report template. Available at: <http://som.flinders.edu.au/FUSA/SACHRU/PDF/EvalReportTemplate.pdf> (accessed Sep 2007).
- 15 Sanders D, Labonte R, Baum F, Chopra M. Making research matter: a civil society perspective on health research. *Bull World Health Organ* 2004; 82: 757-63.

(Received 12/12/06, revised 21/05/07, accepted 31/05/07) □